Università degli Studi di Pisa
Scuola di Dottorato Galileo Galilei – Dottorato in Fisica

Study of Flavour Changing Neutral Currents in top quark decays with the CMS detector

Ph.D. Thesis

December 2007

Ph.D. student:
Dott. Leonardo Benucci

Advisor:
Prof. Guido Tonelli

Preface

This work contains an estimation of the observability of the Flavour Changing Neutral Current in top decays with the CMS experiment at the future Large Hadron Collider.

The first realistic estimates for the experiment sensitivity to the non-Standard Model decays $t \to Zq$ and $t \to \gamma q$ (where q represents c or u quarks) have been addressed, exploiting the photon and the leptonic decays of the Z boson. The task has been accomplished through the use of a full simulation of sub-detectors and reconstruction chain.

The work is organized as follows.

The first two chapters are a self-contained introductory part. Chapter 1 introduces to the world of the top quark. Starting from its observed features, the nature of this intriguing object is outlined, emphasizing how much the next future experiments can highlight this nature. General topics about top production mechanisms and decay are summarized, along with some insights on the experimental challenges that will be coped by the work. Then a separate section is devoted to Flavour Changing Neutral Current (FCNC): the general theoretical framework is reminded and the possible anomalous coupling manifestations are sketched. The most important message is that, while expected branching ratios for FCNC in the Standard Model are eagerly low ($10^{-12} \div 10^{-14}$), thus impossible to be detected, there are extensions in new physics that may enhance these rates. Peculiar configurations of the Double Higgs or Minimal Supersymmetry models, R-parity violation, as well as left-right asymmetry, techni-color and quark singlet models could lead to branching ratios up to $10^{-4} \div 10^{-5}$, that will be demonstrate to be accessible by this analysis. Therefore, any experimental evidence for a top quark FCNC interaction would signal the existence of physics beyond the Standard Model. In some cases, looking for anomalous decays in the top sector provides a tool to constrain parameters of new models, that are difficult to access by other searches. Finally, an overview of the current limits on the tVq couplings, from both direct searches or indirect limit, is presented. Results from other collaborations at the LHC experiments are mentioned, setting the scene for the new results of the work.

Chapter 2 is a review of the CMS experiment. First the accelerator and the environment

in which it will operate are presented; then motivations for the peculiar CMS design are explained. A general view of the experiment sub-systems and trigger system given.
A description of the algorithms CMS have designed to reconstruct the physical objects (electrons and photons, muons, jets and missing energy, jets from b quark) follows, as an important prerequisite to the analysis phase. For each objects, both the on-line and off-line reconstruction strategies are outlined. Some specifications about the software framework adopted in the simulation and reconstruction follow, with a general description of operations it performs.

The core of the analysis work is contained in Chapter 3. Once the distinctive features of the two addressed signals ($t \to Zq$ and $t \to \gamma q$) are identified, a strategy is assessed to reveal these features on top of the Standard Model background. The background sources having some relevance in this analysis phase are presented, by reporting their cross sections from theoretical calculations and discussing the impact on the signal detection. Next section describes some of the tools that are common to both the analyses, specifying the general reconstruction choices performed and introducing the variables that will be involved in the selection procedure. The analysis of the two FCNC signals are based on a set of cuts performed on Monte Carlo samples, where each cut is aimed to reduce at most the background without depleting too much the efficiency of the signal. Optimization of cuts is driven by maximization of a significance region, whose meaning will be explained in the subsequent chapter. For each analyses, all details are given from the pre-selections of the final state particles (electrons and muons, photons and jets) to the reconstruction of the higher level objects, as W and Z boson and finally the top quark. Through the devised strategy, the background suppression is successfully obtained, the combinatorial can be easily managed and the signal from anomalous top decay can actually emerge.

The selection efficiencies for the signal and background are considered in Chapter 4, that is aimed to estimate the sensitivities of the experiment. Firstly, a full set of systematic effects, originating from detector or theoretical uncertainties, are added to the simulation and their impact on the analysis is established. Though both the analyses have been optimized for the low luminosity phase of LHC, here a step forward is made and a test of how they behave at very high luminosity is performed, by artificially increasing the pile-up in the simulation. Then a closer look is given to the relevant background: assuming that Monte Carlo samples have several limitations and cannot reproduce the reality in all the details, efforts are performed to find a specific control region for both analyses, that can be addressed in the future samples, thus allowing to measure directly the background when data will be available. Calculations and rescaling of the control region to the signal one are documented, from where a more robust estimation of the background is deduced. The tricky issue of which selection significance is to be adopted is then addressed, and solved by adopting an approach that maximizes both the confidence level and the discovery reach. This set the scene to determination of the minimum branching ratio for top anomalous decays that CMS can detect.

Results are presented for several background levels and different experimental conditions, then extrapolated to the highest attainable luminosities, where this topic could be inquired with the best results. Comparison with current limits are then showed.

In the conclusion (Chapter 5) the whole work is summarized and original approaches are underlined. It is emphasized that, thanks to an analysis like the one presented (along with similar studies that can be inspired by its result), at LHC the door is open for new results and possible discoveries.

The appendix contains some interesting topics that are linked to the work, but not essential for its development. They are the role of top mass in EW precision physics, a more detailed description of CMS sub-detectors and their performance, an outline of perspectives for the LHC luminosity upgrade.

Contents

Chapter 1

Top quark at LHC and Flavour Changing Neutral Currents

Though the existence of the top quark was predicted in the late 70s, its search has been a twenty years-long challenge. Top quark was discovered at Fermilab in the 1995-96 [1], completing the three-family structure of the Standard Model (SM) and opening up the new field of top quark physics. Since the beginning of the study phase, this object has appeared as a very special one.

Top quark is distinguished by a large mass (about 35 times larger than that of the lighter quark), that is intriguingly close to the scale of electroweak (EW) symmetry breaking. In addition, top Yukawa coupling is surprisingly close to one. These unique properties suggest that studies of the top quark may provide an excellent probe of electroweak symmetry breaking and could be the key for revealing new physics. Some of the questions that top physics could answer are indeed very fundamental ones:

- is the top quark mass generated by the Higgs mechanism as the SM predicts? How is it related to the top-Higgs Yukawa coupling?
- alternatively, does the top quark play a more fundamental role in the EW symmetry breaking mechanism?
- if new particles lighter than the top quark exist, does the top quark decay into them?
- could non-SM physics manifest itself in non-standard couplings of the top quark, showing anomalies in top quark production and decays?

If some new physics exists, its effect should show up very clearly on top of the precise SM prediction for this particle. Top quark physics are mainly focused on perform such precise measurements and reveal these effects.

Several properties of the top quark have already been examined at the $p\bar{p}$ collider Tevatron at Fermilab (up to now the only place where top quark is directly produced), $e^{\pm}p$ collider

HERA at DESY and e^+e^- collider LEP at CERN (where just indirect measurements have been possible via loop effects).

The two Tevatron experiments, CDF and DØ, have gained a wide experience in top studies and publications are a large amount today. Two different ages mark the life of Tevatron: the Run-I (lasted until 1996), where hadrons collided at $\sqrt{s} = 1.8$ TeV (and a luminosity of $\sim 200\,\mathrm{pb}^{-1}$ was integrated) and the Run-II, started in 2001, with an increased energy $\sqrt{s} = 1.96$ TeV and an upgraded luminosity. Run-II is now in full swing, and recent analyses may profit of the larger integrated luminosity (3.2 fb^{-1} delivered, 2.7 fb^{-1} on tape in November 2007) and the lessons learned from Run-I.

The value of top mass has been measured in all decay channels (di-leptonic, lepton+jets, all-hadronic, see Sec. 1.2), exploiting different analysis techniques. This important parameter is regularly updated: relevant bibliography[1] can be found in Ref. [2] for the Run-I and Ref. [3] for the Run-II. Also the top production cross section (along with its kinematical properties, [4]) has been extracted from all channels (see for example Ref. [5] for Run-I and Ref. [6] for Run-II), including challenging final states with fully hadronic [7] and τ decays [8]. Among top standard decays, precise measurements have been performed, *e.g.* the W boson polarization (with the first significant direct constraint for the V+A contribution in top decay), performed both in Run-I [9] and in Run-II [10], the first measurement of the ratio $BR(t \to Wb)/BR(t \to Wq)$ and limits on the CKM Element V_{tb} [11], as well as bounds on $t\bar{t}$ spin correlations [12]. All these observations reported only SM effects, constraining many new physics parameters. New physics has been searched directly also, looking for flavour changing neutral current decays [13], top decays in charged Higgs boson (in a b quark and τ lepton) [14] and new particles decaying to a $t\bar{t}$ pair [15].

The production of a single-top (outlined in Sec. 1.1) is one of the most challenging Standard Model signal at Tevatron, because production cross section is less than a half of the $t\bar{t}$ one, the signature is less distinctive and background sizable. Several signal extraction techniques have been exploited in both runs [16] and upper limits are set. A quite clear indication of single-top production has been actually provided by DØ at the end of 2006 [17] and has been recently confirmed by CDF [18].

The HERA collider as well has a center-of-mass energy allowing the electro-weak production of a single-top. The two experiments ZEUS and H1 analyzed inclusive data with integrated luminosity exceeding $100\,\mathrm{pb}^{-1}$, focusing searches on a not-SM production via neutral currents (FCNC), the only way in which a top could be produced at such machine [19].

Despite the very important reaches and limits of these measurements, most of them suffer from the small sample of top quarks collected. It is here that the Large Hadron Collider (LHC), near to be lighted-on, comes into the game.

Whereas the center-of-mass energy at the collision at LHC is seven times higher than the Tevatron, the cross section $\sigma(t\bar{t})$ for $t\bar{t}$ production (as Sec. 1.1.3 will explain) is more

[1] In most of references of this paragraph, only the more recent bibliography is mentioned.

than a hundred times larger. This large cross sections, combined with the significantly enhanced single-top production, implies that during the stable low-luminosity run ($\mathcal{L} = 2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$) a rate about 4 top per second will be reached. About 800 k events will be likely obtained after 1 fb^{-1}, then several millions of events can be accumulated each year.

Therefore, LHC can well be renamed as a "top factory". Having a so powerful machine at hand, several top quark properties will be examined with a much better precision and entirely new measurements can be contemplated. Some of items that will be accessible, thanks to a so large statistic, are listed below:

- precision measurements will be possible for a lot of variables in the production and the decay phase (as electric charge, spin polarization, W polarization, decays and coupling etc.), providing several handles on new physics. A good evaluation of $\sigma(t\bar{t})$ is possible and it will be extremely useful for investigations of scenarios beyond Standard Model, where $t\bar{t}$ is often the main background;

- as explained in Sec. A.1, radiative corrections in the EW symmetry breaking model are proportional to m_t^2. Thus a precise top mass (and W mass) determination is important to scrutinize the Standard Model (SM) and provide mass constraint to the Higgs boson;

- the enormous mass of top (compared to the others quarks and lepton) and its large coupling with the Higgs field suggest that top production and decay is a perfect place where new physics could manifest. As an example, in many scenarios beyond SM heavy particles decay into top quarks, both $t\bar{t}$ pair and single-top quark.

In addition, the top pair production is a valuable tool to *in situ* calibration, in the early commissioning stage, of LHC detectors involved in top physics. The large cross section and signal/background ratio for the lepton+jets $t\bar{t}$ channel (see Sec. 1.2) allow high purity sample with large statistic to be produced in a short time period. Understanding experimental signature for top events involves most parts of the detectors and is essential in claiming all potential discoveries.

This chapter is intended to present the most important features of top quark, focusing on the specific research field addressed in this work. Section 1.1 offers an overview of the relevant properties of this particle (mass, width and decays) from a theoretical point of view, along with a phenomenology of its production and decay processes, mostly oriented to the future hadron collider. Section 1.2 explains how experimentalists try to find the top quark and identify its features. Some general concepts about experimental issues are given (that will be regularly used in the rest of the work) and a window toward LHC perspectives is always kept open. Section 1.3 is devoted to a specific item of top quark physics, namely the Flavour Changing Neutral Currents (FCNC) couplings. After a review of the proposed theoretical framework (that assigns to FCNC the role of messengers of new physics), the set of processes where these couplings can manifest are listed and explained. Then the witness is

given to current and recent machines, with a summary of the results drawn from the searches of such processes. Once again, the strong improvements that LHC is supposed to reach are underlined, presenting some expectations from simulation studies in this field, and setting the scene for the venue of the present work.

1.1 Top quark properties

According to the SM, the top quark is a spin-1/2 and charge-2/3 fermion, transforming as a colour triplet under the group $SU(3)$ of the strong interactions and as the weak-isospin partner of the bottom quark. A recent DØ measure [20] swept away the doubt that the observed top would be a $\pm 4/3$ charged particle, decaying as $t \to W^+\bar{b}$ or $t \to W^-b$.

It is worth to observe that the analysis of EW observables in Z^0 decays [21], performed well before the top discovery, required the existence of a $T_3 = 1/2$, charge $+2/3$ fermion, with a mass in the range of $\approx 170\,\mathrm{GeV}/c^2$. As it will be reported in Sec. 1.1.1, direct Tevatron measurements agree very well with this estimates. Also measurements of the total cross section at the Tevatron, when compared with the theoretical estimates, are consistent with the production of a spin-1/2 and colour-triplet particle.

Some basic properties of top quark are presented in the following, as deduced from the Standard Model expectations (see e.g. Ref. [22]) and extracted from up-to-date experimental results. Thanks to the lack of a top spectroscopy (Sec. 1.1.2), predictions for almost all top quark interactions can be evaluated using perturbation theory, thus avoiding uncertainties due to fragmentation processes.

1.1.1 Top mass

In addition to its quantum numbers, the two most fundamental properties of the top quark are its mass m_t and width Γ_t.

Here m_t is intended to be the *pole* top mass of the Breit-Wigner shape. In the SM, m_t is related to the top Yukawa coupling through the Fermi constant G_F as:

$$y_t(\mu) = 2^{3/4} G_F^{1/2} m_t \left(1 + \delta_t(\mu)\right),$$

where $\delta_t(\mu)$ accounts for radiative corrections at a μ scale. The top pole mass, like any quark mass, is defined up to an intrinsic ambiguity of order $\Lambda_{QCD} < 200\,\mathrm{MeV}$ [22] because the top production and decay process is complicated by hadronization effects which connect the b quark from top decay to other quarks involved in the original scattering.

Current top mass estimate

The current value of top pole mass is obtained combining measurements from the CDF and DØ experiments during Run-I with the most recent Run-II ones. Taking correlated errors

properly into account and assuming Gaussian systematic uncertainties, the resulting world average mass of the top quark is today (Spring 2007):

$$m_t = 170.9 \pm 1.1\,(\text{stat}) \pm 1.5\,(\text{syst})\,\text{GeV}/c^2\,, \tag{1.1}$$

corresponding to a $1.8\,\text{GeV}/c^2$ total error, *i.e.* 1.1% precision [23]. This latest results exploits an amount of $\sim 1\,\text{fb}^{-1}$ statistic on tape at Tevatron.

The impact of m_t on Higgs mass is a key issue in next-to-come electroweak precision physics. Nevertheless, the topic is not directly related to the present work, thus an outline is in the appendix, Sec. A.1.

1.1.2 Top quark couplings: decay channels and mass width

One of the consequences of the large top quark mass is a large value for its mass width, well exceeding the QCD hadronization scale ($\Lambda_{\text{QCD}} \sim 0.2\,\text{GeV}$). This implies that the top quark lifetime ($\tau(t) \simeq 4.6 \cdot 10^{-25}\,\text{s}$) is small compared to the time scale for hadronization ($\sim 3 \cdot 10^{-24}\,\text{s}$): hence the top quark exists only as a free quark, it decays before any hadronization occurs and top hadrons (mesons or baryons) are definitely ruled out.

According to the the Standard Model CKM matrix, the only relevant coupling of t is with b quark because $|V_{tb}| = 0.999100^{+0.000034}_{-0.000004}$, $|V_{ts}| = 41.61^{+0.12}_{-0.78} \cdot 10^{-3}$ and $|V_{td}| = 8.14^{+0.32}_{-0.64} \cdot 10^{-3}$, as resulting form the global fit reported in Ref. [24] (that assumes 3 families and unitarity). Therefore, $t \to bW$ is by far the dominant decay mode ($BR > 99.7\%$) and in fact the only observed up to now. Other qW decays widths are very small ($BR(t \to sW) = 1.23 \div 1.76 \cdot 10^{-3}$, $BR(t \to dW) = 0.16 \div 1.71 \cdot 10^{-4}$). The b quark fragmentation functions control the formation of hadrons in which top decays: $BR(t \to B^- X) \simeq BR(t \to B^0 X) \simeq 40\%$, $BR(t \to B_s X) \simeq BR(t \to \Lambda_b X) \simeq 10\%$.

The on-shell decay width Γ_t is usually expressed in units of the lowest order decay width with M_W and m_b set to zero and $|V_{tb}|$ set to 1:

$$\Gamma_0 = \frac{G_F m_t^3}{8\pi\sqrt{2}} = 1.76\ \text{GeV}\,.$$

Incorporating M_W the leading order result reads

$$\Gamma_{\text{LO}}(t \to bW)/|V_{tb}|^2 = \Gamma_0 \left(1 - 3\frac{M_W^4}{m_t^4} + 2\frac{M_W^6}{m_t^6}\right) = 0.885\,\Gamma_0 = 1.56\ \text{GeV}\,.$$

This lowers to 1.42 GeV when radiative QCD and EW corrections are included. Theoretical uncertainties on this value are below 1%.

Beyond the three level, several other top decays are possible, although with an extremely low branching ratio, such as: $t \to bWZ$, almost 'closed' because $m_t \sim m_b + M_W + M_Z$ and thus very sensitive to m_t ($BR = (1 \div 10) \cdot 10^{-7}$ with $m_t = 179 \div 169\,\text{GeV}/c^2$) [25]; $t \to cW^+W^-$, that is further suppressed ($BR \sim 10^{-13}$) because of the GIM rule in the $V_{tj}V_{jc}^*$

coupling; $t \to bWH$, having $BR \sim 7 \cdot 10^{-8}$ provided that $m_H > 100\,\mathrm{GeV}/c^2$ [26]. New physics, as the existence of a b' belonging to a 4^{th} family with a mass below few hundreds GeV, may lead to detectable rates.

1.1.3 Top quark couplings: production processes

The large top quark mass ensures that top production is a short-distance process, and that the perturbative expansion, given by a series in powers of the small parameter $\alpha_S(m_t) \sim 0.1$, converges rapidly. Top quark has two ways to spread out from hadrons collisions: $t\bar{t}$ associated production (occurring via QCD-related processes) and single-top production trough electroweak mechanisms. A brief review of the current knowledge is presented below [22].

Top pair production

Associated production of a top-anti-top pair happens via the $gg \to t\bar{t}$ or $q\bar{q} \to t\bar{t}$ channels, as shown in Figure 1.1.

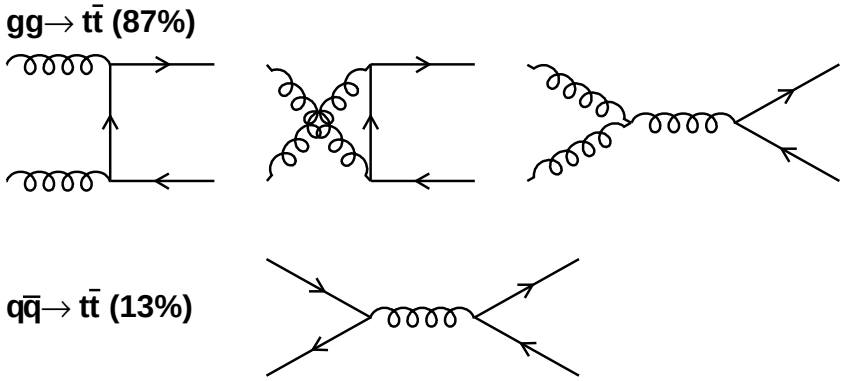

Figure 1.1: Production of a $t\bar{t}$ pair may occur both via gg fusion (upper diagrams) and $q\bar{q}$ annihilation. At LHC, the former is expected to contribute for about a $87 \div 90\%$ of cross section, while at Tevatron quark annihilation is dominant.

At LHC energies ($\sqrt{s} = 14$ TeV), the largest of the proton parton distribution functions (PDFs) is the gluon density, making $gg \to t\bar{t}$ the principal top production mechanism ($\sim 90\%$). At Tevatron quarks annihilation is the dominant process and weights of two production mechanism are roughly reversed.

Total $t\bar{t}$ production cross sections calculation is available at next-to-leading-order (NLO, $\mathcal{O}(\alpha_S^3)$) [27]. Theoretical progresses over the last years has led to the resummation of Sudakov-type logarithms [28] which appear at all orders in the perturbative expansion for the total cross sections. More recently, the accuracy of these resummations has been extended to the next-to-leading logarithmic (NLL) [29]. While the inclusion of the higher-order terms does

Table 1.1: Resummation contributions to the total $t\bar{t}$ cross sections ($m_t = 175\,\mathrm{GeV}/c^2$) in pb. PDF set is MRST. A indicates different NNL models for NLL calculation. All results are evaluated at $\sqrt{s} = 14\,\mathrm{TeV}$ [22].

		NLL resummed, A=2		NLL resummed, A=0	
$\mu_R = \mu_R$	NLO	$\mathcal{O}(\alpha_S^{\geq 4})$	NLO+NLL	$\mathcal{O}(\alpha_S^{\geq 4})$	NLO+NLL
$m_t/2$	890	-7	883	-12	878
m_t	796	29	825	63	859
$2m_t$	705	77	782	148	853

Table 1.2: Total $t\bar{t}$ cross sections ($m_t = 175\,\mathrm{GeV}/c^2$) in pb, evaluated at NLO+NLL ($A = 0$) [22].

PDF	$\mu = m_t/2$	$\mu = m_t$	$\mu = 2m_t$
MRST	877	859	853
MRST $g\uparrow$	881	862	857
MRST $g\downarrow$	876	858	852
MRST $\alpha_S\downarrow$	796	781	777
MRST $\alpha_S\uparrow$	964	942	934
CTEQ5M	904	886	881
CTEQ5HJ	905	886	881

not affect significantly the total production rate, it enforces the theoretical predictions under changes in the renormalization and factorization scales, hence improving the predictive power. Being a pure QCD process, uncertainties in $\sigma(t\bar{t})$ cross section comes from renormalization and factorization scale and PDF.

To evaluate the sensitivity to renormalization (μ_R) and factorization (μ_F) scale, the two parameters are typically varied over the range $\mu_0/2 < \mu < 2\mu_0$, keeping $\mu_R = \mu_F = \mu_0 = m_t$. A detailed breakdown of the NLO $\mathcal{O}(\alpha_S^3)$ and higher-order $\mathcal{O}(\alpha_S^{\geq 3})$ contributions, as a function of the scale and of the value of the parameter A (that is related to a specific structure of higher order, NNLO), is given in Table 1.1. All results are evaluated at $\sqrt{s} = 14\,\mathrm{TeV}$. After inclusion of NLL corrections, the scale uncertainty is significantly reduced: it amounts to a $\pm 6\%$ variation in the most conservative case.

The parton distribution functions dependence is at the level of $\pm 10\%$ and it is given in detail in Table 1.2 for $m_t = 175\,\mathrm{GeV}/c^2$, choosing some different PDFs.

Combining these predictions, taking a central value for $m_t = 175\,\mathrm{GeV}/c^2$ and $\mu = m_t$, the $t\bar{t}$ cross section is expected to be:

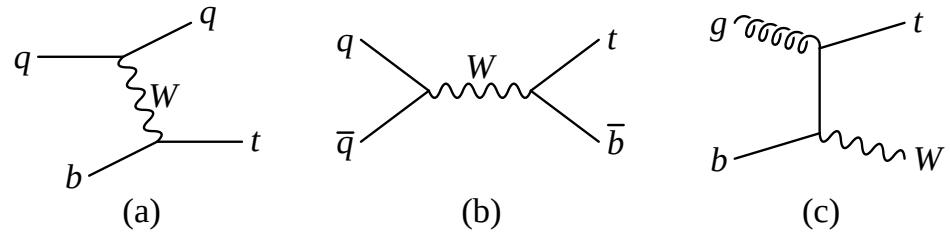

Figure 1.2: Feynman diagrams for single-top quark production in hadron collisions: (a) t-channel process; (b) s-channel process; (c) associated production (only one of the two diagrams for this process is shown).

$$\sigma(t\bar{t}) = (833 \pm 83\,(\mathrm{PDF}) \pm 50\,(\mathrm{stat}))\,\mathrm{pb}\ . \tag{1.2}$$

As Tab. 1.2 suggests, $t\bar{t}$ production cross section is a function of the top mass, roughly proportional to $1/m_t^2$.

Single-top production

Top quarks can be produced at colliders without an associated anti-top. The electro-weak processes that make this possible are listed below and Feynman diagrams showed in Fig. 1.2. The total cross sections for the three single-top quark production processes in the SM, calculated at LHC energy (with the PDF CTEQ6M and $m_t = 175\,\mathrm{GeV}/c^2$) are summarized in Table 1.3, along with their theoretical uncertainty [22]. Other references to intensive studies on the topic are in Ref. [30].

t-channel :

The dominant process involves a space-like W boson ($q_W^2 > 0$) [31], striking a b quark in the proton sea and promoting it to a top quark. This process is also referred to as W-gluon fusion, because the b quark ultimately arises from a gluon splitting to $b\bar{b}$. Since by definition is $|q^2| \leq M_W^2$, the final state light quark tends to be emitted at small angles, *i.e.* high rapidities. This characteristic proves to be useful when isolating this signal from backgrounds (see Sec. 1.2). The b distribution function in the proton sea arises from the splitting of virtual gluons into nearly-collinear $b\bar{b}$ pairs: therefore it is implicit that there is a $\bar{b}$ in the final state, which accompanies the top quark and the light quark. The final-state $\bar{b}$ tends to reside at small p_T, so it is usually unobservable. At LHC, the t-channel process has a cross section nearly one third as large as the cross section for top quark pairs. It has been calculated at NLO, including corrections associated with the light quark, gluon radiation and $b\bar{b}$ pair (non-collinear) splitting. Since LHC is a pp collider, the cross section for t and $\bar{t}$ are generally different: for this channel, it amounts to 152.6 pb for t and 90.0 pb for $\bar{t}$ [31]. The central value for the cross section is obtained by setting the μ_0 scale of the b distribution function equal to $\mu_0^2 = -q^2 + m_t^2$. The uncertainty in the NLO cross section due to the variation of the

factorization scale (between one half and twice its central value) is 4%. In addition, an uncertainty in the top quark mass of 2 GeV yields a total of 2% in the cross section and the error due to the parton distribution functions is estimated to be 10% [32]. Combining all sources in quadrature, the total theoretical uncertainty is presently 11% in the total cross section, rising to 15% for the Wbj (where the jets from the low p_T $\bar{b}$-quark cannot be observed);

s-channel (or W^* process):

The process producing a single-top associated with a b quark can be viewed by simply rotating the t-channel diagram, such that the virtual W boson becomes time-like ($q_W^2 \geq (m_t + m_b)^2$) [33]. The s-channel cross section is much smaller than the t-channel one and amounts to 6.55 pb for t and 4.07 pb for $\bar{t}$ (evaluated at NLO). Nevertheless, this process has the advantage that the quark and antiquark distribution functions are relatively well known, so the uncertainty from the parton distribution functions is small (around 4%). The parton luminosity can be further constrained by measuring the Drell-Yan process $q\bar{q} \to W^* \to l\bar{\nu}$, which has the identical initial state. Factorization and renormalization scales varying between $\sqrt{q^2}/2$ and $2\sqrt{q^2}$ yields a 2% uncertainty. A large sensitivity of cross section to top quark mass (5% for 2 GeV) has to be added and a total uncertainty of 7% is obtained. This is much less than the present theoretical uncertainty in the t-channel cross section;

Wt **associated production** :

A single-top quark may also be produced via the weak interaction in association with a real W boson ($q_W^2 = M_W^2$) [34]. One of the initial partons is a b quark in the proton sea, as in the t-channel process. Unlike the t-channel process, this process scales like $1/s$: this, combined with the higher values of x needed to produce both a top quark and a W boson, leads to a cross section for associated production at LHC which is about a factor of five less than of the the t-channel process, despite the fact that it is of order $\alpha_S \alpha_W$ rather than α_W^2. Here cross sections for t and $\bar{t}$ are the same, giving 64 pb for $W^+\bar{t} + W^-t$ (LO, with a subset of the NLO corrections included). The uncertainty in the cross section due to the $\sqrt{s}/2 < \mu < 2\sqrt{s}$ variation is 15%, to be added to those due to parton distribution functions (10%) and top quark mass (4%). Combining all theoretical uncertainties in quadrature delivers a total uncertainty at present of 18%, the largest of the three single-top processes.

Table 1.3 collects the cross section of the three production systems at LHC, with their theoretical uncertainties.

It is worth to note that cross sections for single-top quark processes are proportional to $|V_{tb}|^2$: therefore, these processes provide the only known way to directly measure V_{tb} at hadron colliders.

Table 1.3: Total cross sections (pb) for single-top quark production at the LHC, for m_t=175 GeV/c^2 [22]. Each cross sections include the values for the t and $\bar{t}$ production. The uncertainties are due to variation of the factorization and renormalization scales, parton distribution functions and top quark mass (2 GeV).

process:	t-channel	s-channel	Wt
σ(pb)	245	10.2	60
$\Delta\sigma$(pb)	27	0.7	11

While the Tevatron Run-I failed in searching for single-top production, recent analysis of Run-II data showed first evidences of this phenomenon [17, 18] and confirmations are coming as more and more luminosity is integrated. On the other hand, the much larger cross sections at LHC should lead to observe the first Wt process and to measure single-top process observables with a wide statistic and an unprecedent accuracy.

1.2 Top quark observability at LHC and experimental issues

As outlined in the previous section, the expectation for LHC to become the first 'top factory', starting from its very initial phase, is based on a firm ground. During the stable run with $\mathcal{L} = 2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ more than 8 million $t\bar{t}$ pairs and 3 million single-t will be produced per year per experiment, in fact opening a new era in our top physics knowledge. To profit of the full potential of such huge production, the proper reconstruction and analysis technique have to be developed and optimized. This short section is devoted to a quick review of general experimental issues, that will be considered a background for the rest of this work. The focus is explicitly tuned on the two LHC experiments that are supposed to do good top physics: ATLAS and CMS.

1.2.1 The challenges of top reconstruction

A top quark has to be reconstructed by searching for the final state of its decay products. Since the only decay channel observed up to now is $t \to bW$, top identification is related to W boson decay mode and b quark hadronization.

W boson decay mode : In approximately 67.6% events, a W boson decays hadronically via $W \to jj$, leading to two or more jets. The branching ratios for leptonic decay with a muon or an electron is about 22% and 10.8% is left for $W \to \tau\nu$. Hadronic and tau decays are generally difficult to extract cleanly above the large QCD background: therefore most of the analyses involving top quarks (including the present one) have at least one of them with W boson decay ($W \to l\nu$, $l = e, \mu$). The transverse momenta of

the leptons coming from W tend to be much higher than those from other sources, thus this quantity, combined with the large missing energy due to the escaping neutrino(s) (here indicated by E_T^{miss}), provides a large suppression against multi-jet backgrounds;

b quark hadronization : In general, jets from background have a steeply falling E_T spectrum, an angular distribution peaked at small angles with respect to the beam and contain b and c quarks at the few percent level. On the contrary, for the top signal all the final states are supposed to contain at least one b-jet and the jets are rather energetic, since they come from the decay of a massive object. An important experimental tool for selecting clean top quark samples is the ability to identify b-jets (*b-tagging*), *i.e.* jets coming from the decay of an hadron containing a b quark, as well as selecting the most energetic and central kinematic region. Techniques for b-tagging, using secondary vertexes, semi-leptonic b-decays and other characteristics of b-jets, have been extensively studied. Both ATLAS and CMS expect to achieve, for a b-tagging efficiency of 60%, a rejection of at least 100 against 'prompt' jets (those containing no long-lived particles) at low luminosity. At high luminosity, a rejection factor of around 100 can be obtained with a somewhat reduced b-tagging efficiency of typically 50%.

Once a physics analysis reaches a good identification of W boson and b-jet, together with other objects supposed to form the final state, the presence of the top quark can be identified by means of invariant mass plot or more refined techniques. Nevertheless, the task to reconstruct the originary top is not accomplished yet. A set of physical effects may be involved in some of the steps from top production to the observed final state: if not properly accounted for, they distort measurements in a systematic way, so they are usually referred to as 'systematic uncertainties'. In top quark reconstruction at a hadron collider, several effects may contribute to systematic uncertainties:

- in the parton-parton collisions at $\sqrt{s} = 14\,\text{TeV}$, gluons are regularly radiated both from the initial parton (ISR) and from the final ones (FSR). This radiation decreases the effective parton energy in the hard interaction, from one side, and increases the number of jets produced in the decay of top quarks, from the other;

- most of theoretical predictions rely on the choice for the Q^2 hard process scale, whose value is determined by the hard process under study. It directly enters in the parameterization of PDFs and cross sections but the dependence of the observables from it is unphysical. The sensitivity of the predicted observables to the Q^2 choice is expected to decrease with the increasing order in which the calculation is performed;

- details of the jets structure (*e.g.* their fragmentation function and their shapes), may influence the experimental determination of the jet energy scales, as well as the efficiency with which b-jets will be tagged;

- instantaneous luminosity delivered by the machine may fluctuate of some percent, inducing a corresponding effect on the rate of each process;

- the parton distribution functions of interacting particles, that describes the probability density for partons undergoing hard scattering at the hard process scale Q^2 and taking a certain fraction x of the total particle momentum. Since the Q^2 evolution of these functions can be calculated perturbatively in the framework of QCD, PDFs measurements can be cross–checked using heterogeneous DIS from other experiments, Drell-Yan and jet data. With these semi-empirical models, predictivity is achieved for points where no direct measurements are available yet, for example in a large region of the (x, Q^2) space for pp interactions at the LHC energy. Various approaches are currently available to quote the PDFs of the proton, which propose different solutions for what concerns the functional form, the theoretical scheme, the order of the QCD global analysis (including possible QED corrections), and the samples of data retained in the fits. The CTEQ [35] and MRST [36] PDFs, including Tevatron jet data in the fits, seem to be well suited for use in Monte Carlo simulations for the LHC;

- many analyses make use of theoretical calculations as input values for some algorithms, as well as for events generation with simulation tools. Previous section has mentioned how sensitive many observables are to the top mass: therefore, robustness of experimental results toward input m_t value should always be checked.

The hard scattering event is accompanied by two processes: the so-called underlying event (UE), which identifies all the remnant activity from the same proton-proton interaction and whose definition often includes ISR as well, and the 'pile-up', composed by other interactions ('minimum-bias') superimposed on the signal events, occurring in the same bunch crossing. Minimum bias events are dominant at LHC but, as there is little detailed theoretical understanding about them, event generators must rely on present data.

1.2.2 Observation of top pairs

The topology of top-anti-top production is naturally driven by the top quark final state. Three classes of events with different branching ratios (BRs) are usually searched for:

semi-leptonic (e/μ+ light-jets + b-jets) ($BR \sim 6/9 \times 2/9 \times 2 \sim 29.6\%$). The signal is easily triggered by the hard lepton from the leptonic W decay. The top mass can be reconstructed from the hadronic decay side exploiting the $M(jjb)$ invariant mass, while the leptonic top decay can be identified in the transverse plane (with quadratic ambiguity) by imposing $E_T^{miss} = p_T(\nu)$ and $M(l\nu) = M_W$;

leptonic (di-e/μ + b-jets) ($BR \sim 2/9 \times 2/9 \sim 4.9\%$). The signal is easily triggered by the hard di-lepton from the two W decays. Due to the presence of two neutrinos in

the final state, the mass measurement relays on the Monte Carlo studies for the angle between the b quark and the lepton, as well as invariant mass of two leptons. The pairing between two objects is made by minimizing the invariant mass $M(lb)^2$ and is correct in about 85% of cases;

hadronic (multi-jets) ($BR \sim 6/9 \times 6/9 \sim 44.4\%$). Triggering the hadronic channel is far from trivial, since only very high jet energy thresholds give reasonable QCD rejection. Neverthless, a reconstruction of the complete event is still possible, through an extensive use of event-shape variables

Due to the very large samples of top quarks which will be produced at the LHC, measurements of the total cross section $\sigma(t\bar{t})$ will be limited by the uncertainty of the integrated luminosity determination, which is currently estimated to be 5%-10%. The cross section relative to some other hard process, such as Z production, should be measured more precisely.

1.2.3 Observation of single-top process

Differently from most of CDF and DØ searches (that generally look for an inclusive single-top production), LHC experiments want to develop efforts to separate the three production channels, in order to enlighten the different physics hidden in them. A presentation of their features may help to identify the specific selection strategies:

t-channel : The most striking feature of the final state for this process is the presence of a forward light-jet from the "spectator" quark, i.e. the one recoiling against the W. In addition, the $\bar{b}(b)$ quark associated to the $t(\bar{t})$ quark tends to be produced at low p_T and very small angle, resulting outside of the detector acceptance in most cases. Therefore, the dominant final state is tj (where the jet is at high rapidity) and the typical selection requires exactly two jets with only one tagged as b-jet. The $Wb\bar{b}j$ final state (accounting for roughly 40% of the events when $p_T(j) > 20\,\mathrm{GeV}/c$) is not exploited usually, because of the large background from $t\bar{t}$. The cross section for the inclusive process tj is reduced to 164 pb ($\pm 10\%$) when $p_T < 20\,\mathrm{GeV}/c$;

s-channel : Here the presence of a top quark associated with a hard b-jet calls for the requirement of exactly two high p_T jets, both tagged as b-jets;

tW-channel : This process yields two W bosons, the one produced in association with the top and the other coming from its decay.

More refined methods based on jet-charge, amount of total transverse energy or improved jet quality requirements, though not described here, have demonstrated to be efficient in future CMS and ATLAS experiments.

1.3 FCNC in top sector

An extensive study of top quark couplings will be possible thanks to the huge LHC statistic. This machine will allow the standard couplings to be scrutinized and to look for couplings with new particles, to observe rare decays for the first time, as well as search for processes forbidden (or highly suppressed) in the Standard Model. Flavour Changing Neutral Currents (FCNC) are one of the most interesting processes in this respect: since they are imputable to non-SM effects, any observation is a signature of new physics.

LHC will permit to discover such effects, if present, or alternatively put much stronger constraints to new physics models. This section offers a general overview of the specific top FCNC mode. The relevant theoretical informations on the most interesting processes are summarized, followed by a brief description of experimental results for current and recent experiments, along with predictions from future ones based on simulations.

1.3.1 Top anomalous interaction: a theoretical framework

In the SM the couplings of the top quark are uniquely fixed by the gauge principle, the structure of generations and the requirement of a lowest dimension interaction Lagrangian. Due to the large top mass, top quark physics looks simple in this renormalizable and unitary quantum field theory: the top quark decay is described by pure V-A weak interactions and only one significant decay channel is present: $t \to bW^+$, other decay channels being very suppressed by small mixing angles. This simplicity makes the top quark a unique place to both constraining the Standard Model and probing possible scenarios for new physics. If anomalous top quark vertexes exist, additional anomalous couplings such as the right-handed vectorial and the left- and right-handed tensorial can also be considered. They will affect top production and decay at high energies, as well as precisely measured quantities with virtual top quark contributions.

Top quark couplings can be parametrized in a model independent way by an effective Lagrangian. After the gauge symmetry breaking, the dimension 5 couplings to one on-shell gauge boson is the sum of four contributions [37]:

$$
\begin{aligned}
\mathcal{L}_{tq}^{g} &= -g_S \sum_{q=u,c,t} \frac{\kappa_{tq}^{g}}{\Lambda} \bar{t}\sigma^{\mu\nu} T^a (f_{tq}^{g} + i h_{tq}^{g} \gamma_5) q G_{\mu\nu}^{a} \quad , \\
\mathcal{L}_{tq}^{W} &= -\frac{g}{\sqrt{2}} \sum_{q=d,s,b} \frac{\kappa_{tq}^{W}}{\Lambda} \bar{t}\sigma^{\mu\nu} (f_{tq}^{W} + i h_{tq}^{W} \gamma_5) q W_{\mu\nu}^{+} \quad , \\
\mathcal{L}_{tq}^{\gamma} &= -e \sum_{q=u,c,t} \frac{\kappa_{tq}^{\gamma}}{\Lambda} \bar{t}\sigma^{\mu\nu} (f_{tq}^{\gamma} + i h_{tq}^{\gamma} \gamma_5) q A_{\mu\nu} \quad , \\
\mathcal{L}_{tq}^{Z} &= -\frac{g}{2\cos\theta_W} \sum_{q=u,c,t} \frac{\kappa_{tq}^{Z}}{\Lambda} \bar{t}\sigma^{\mu\nu} (f_{tq}^{Z} + i h_{tq}^{Z} \gamma_5) q Z_{\mu\nu} \quad , \\
&\quad -\frac{g}{2\cos\theta_W} \sum_{q=u,c,t} \bar{t}\sigma^{\mu} (v_{tq}^{Z} - a_{tq}^{Z} \gamma_5) q Z_{\mu} \quad , \qquad (1.3)
\end{aligned}
$$

plus the hermitian conjugate operators for the flavour changing terms. $G_{\mu\nu}^{a}$ is $\partial_\mu G_\nu^a - \partial_\nu G_\mu^a$ and similarly for the other gauge bosons; e, g_S, g are the electric, strong and electroweak charges. The $\mathcal{L}_{tq}^{Z}$ Lagrangian has contribution from the tensorial and V-A interaction.

The couplings in Eq. 1.3 are usually normalized by taking $\Lambda = 1\,\text{TeV}$. κ is real and positive and f, h are complex numbers satisfying for each term $|f|^2 + |h|^2 = 1$. Renormalization procedures have only a small effects on them. Since in the SM these anomalous couplings are absent at tree-level and occur only at loop level, rates for FCNC processes in the top sector are extremely small [38, 39], because of the strong loop suppression and the high masses of the gauge bosons. Therefore, the top quark plays a unique role compared to the other quarks – for which the expected FCNC transitions are much larger – and *any experimental evidence for a top quark FCNC interaction would signal the existence of new physics.* This can be considered the key sentence that motivates this work.

The fact that a measurement of the top width is not available and the branching ratio $BR(t \to bW)$ is a model dependent quantity makes the present experimental constraints on the top FCNC decays quite weak. Anomalous top interaction can be measured (or constrained) observing either flavour changing charged currents couplings as the standard tWb, or flavour changing neutral ones. At a fundamental level, the latter consist in couplings of the type tVq, where V is a neutral gauge boson ($V = \gamma, Z^0, g$) and q is an up-quark type different from top ($q = u, c$). As it will be shown below, these couplings can be studied in top production and in its decay: both kind of processes contribute to improve the knowledge of anomalous coupling constants, that here is the ultimate goal.

There are different reasons – some *a posteriori* derived – that suggest to drive the focus of the work on the search for FCNC in the decays, thus preferring them to the production processes. Among the relevant motivations:

- decays such as $t \to qZ^0$ and $t \to q\gamma$ have a very clear experimental signature. In other words, many handles are provided to discriminate these modes from the multiform

background processes that LHC will produce. As it will be shown at the end of the work, such decays are also prone to be searched in future LHC upgrades;

- they offer a direct estimate of the strength of the tVq couplings, that can be extracted from the anomalous decays branching ratios in a straightforward way. Thus the typical uncertainties related to the parton content in the initial state (such PDFs and energy scale) are not an issue;
- these analysis will provide tools that can turn useful to explore some other rare top decays, as the three-body modes $t \to qV_iV_j$, that similarly are able to claim for new physics.

Later in this section, some mention will be done for results in FCNC production phenomenology, since the present experimental bounds they delivered are compelling with those from FCNC decays.

1.3.2 FCNC in top quark decays

In the effective Lagrangian description of Eq. 1.3, it is straightforward to calculate the top quark decay rates as a function of the top quark FCNC couplings [40]:

$$\Gamma(t \to qg) = \left(\frac{\kappa^g_{tq}}{\Lambda}\right)^2 \frac{8}{3}\alpha_S m_t^3, \tag{1.4}$$

$$\Gamma(t \to q\gamma) = \left(\frac{\kappa^\gamma_{tq}}{\Lambda}\right)^2 2\alpha m_t^3, \tag{1.5}$$

$$\Gamma(t \to qZ)_\gamma = \left(|v^Z_{tq}|^2 + |a^Z_{tq}|^2\right) \frac{\alpha\, m_t^3}{4M_Z^2 \sin^2 2\theta_W} \left(1 - \frac{M_Z^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_Z^2}{m_t^2}\right), \tag{1.6}$$

$$\Gamma(t \to qZ)_\sigma = \left(\frac{\kappa^Z_{tq}}{\Lambda}\right)^2 \frac{\alpha\, m_t^3}{\sin^2 2\theta_W} \left(1 - \frac{M_Z^2}{m_t^2}\right)^2 \left(2 + \frac{M_Z^2}{m_t^2}\right). \tag{1.7}$$

For comparison, Table 1.4 collects the rare top decay rates normalized to $\kappa^g_{tq} = \kappa^\gamma_{tq} = |v^Z_{tq}|^2 + |a^Z_{tq}|^2 = \kappa^Z_{tq} = 1$, and for the SM, assuming $m_t = 175\,\mathrm{GeV}/c^2$, $\Lambda = 1\,\mathrm{TeV}$, $\alpha(M_Z) = \frac{1}{128.9}$, $\alpha_S = 0.108$. The decays into $q = u, c$ are summed together. In this 'extreme' case with the anomalous couplings equal to one, the top can decay into a gluon, a photon or a Z boson plus a light quark $q = u, c$ at rates similar to the standard bW. Here results have been referred to the update study in Ref. [40].

The complete predictions for the $t \to qV$ branching ratios are: $BR(t \to (c+u)g = 4.6^{+1.1}_{-0.9} \pm 0.2 \pm 0.4^{+2.1}_{-0.7}) \cdot 10^{-12})$, $BR(t \to (c+u)\gamma = (4.6^{+1.2}_{-1.0} \pm 0.2 \pm 0.4^{+1.6}_{-0.5}) \cdot 10^{-14})$, where uncertainties are associated to the top and bottom quark masses, the CKM matrix elements and the renormalization scale. The u quark contributes only for V^2_{ub}/V^2_{cb} to the total rate.

In the Standard Model, flavour changing neutral decays are possible only via loop processes. Two Feynman diagrams for the electroweak emission of a Z/γ are represented in Fig. 1.3; the gluon emission may only occurs from the fermion down-quark side.

Table 1.4: Top quark decay widths and corresponding branching ratios, for the anomalous couplings equal to one and for the SM.

	Top decay mode				
	W^+b	$(c+u)g$	$(c+u)\gamma$	$(c+u)Z_\gamma$	$(c+u)Z_\sigma$
FCNC coupling	–	1	1	1	1
$\Gamma(\kappa=1)$(GeV)	1.56	2.86	0.17	2.91	0.14
BR($\kappa=1$)	0.20	0.37	0.022	0.38	0.018
FCNC coupling	–	$4.8\cdot10^{-5}$	$2.9\cdot10^{-5}$	$9.9\cdot10^{-8}$	0
Γ(SM)(GeV)	1.56	$7.4\cdot10^{-12}$	$7.4\cdot10^{-14}$	$1.6\cdot10^{-14}$	–
BR(SM)	1	$4.6\cdot10^{-12}$	$4.6\cdot10^{-14}$	$\approx10^{-14}$	0

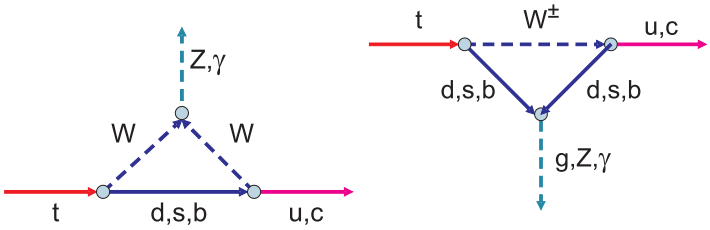

Figure 1.3: Two possible Feynman diagrams for top quark coupling with u, c quark in a Standard Model framework. Only these loops interaction are possible and, due to the GIM rule, they result extremely suppressed.

In particular, in the $t \to cV_i$ transitions the scale of the partial widths $\Gamma(t \to cV_i)$ is set by the b quark mass, and is given by:

$$\Gamma(t \to cV_i) = |V_{bc}|^2 \alpha^2 \alpha_i m_t \left(m_b/M_W\right)^4 \left(1 - m_{V_i}^2/m_t^2\right)^2 ,$$

where α_i is the respective coupling for each gauge boson V_i. It is important to observe that there is a contrast with FCNC transitions in other sectors: in $b \to s\gamma$, for instance, the leading contribution is proportional to m_t^4/M_W^4, thus the GIM mechanism induces an enhancement factor.

In recent years, a wide range of studies addressed FCNC top quark decays in new physics models (starting with papers as Ref. [41]), such as the two Higgs doublet models (2HDM) or the Minimal Supersymmetric Standard Model (MSSM). Their major results, briefly summarized below, are mainly focused on the decays $t \to cV$ (V=γ,g or Z). In many among the mentioned models, an impact also on the $t \to cH$ decay, as well as three-body ones, is expected.

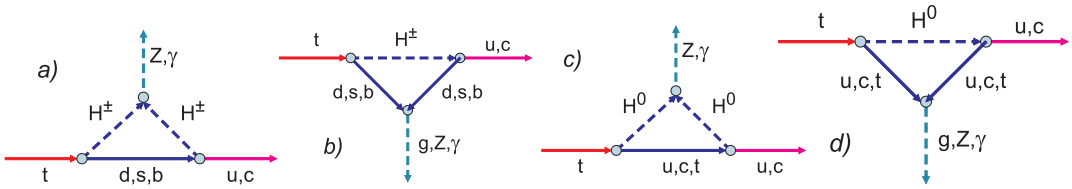

Figure 1.4: With the introduction of 2HDM models, the possibility of additional loops with new gauge bosons opens up. In the I and II models only charged couplings (a) and b)) are possible, while in the model III discrete symmetry preventing charged couplings is removed, thus the amplitude is enhanced by the neutral Higgs bosons exchange (c) and d)).

$t \to Vq$ in 2HDM

The possibility that the electroweak symmetry breaking involves more than one Higgs doublet is well motivated theoretically. There are three possible versions for the 2 Higgs Doublet Model (2HDM), called Type I, II and III 2HDM [42], depending on how the two doublets couple to the fermion sector. All classes have been examined in connection with rare top decays, as they allow to enhance the rates for loop-level and tree processes.

In model I and model II, the up-type and down-type quarks couple to the same scalar doublet and to two different doublets, respectively. Both models are characterized by an *ad hoc* discrete symmetry which include natural flavor conservation, thus forbid tree-level FCNCs [43]. These decay modes are dominated by the one-loop diagrams with a virtual $H^{\pm}$, whose an example is provided by the loop in Fig. 1.4(left). The contribution of charged Higgs loops depend on M_H and $\tan\beta$, but in general is much larger than the W-boson loops since the Yukawa couplings are proportional to fermion masses. Branching fractions for $t \to qg/Z/\gamma$ have few chances to approach a detectability threshold [44]. The loop-induced flavour-changing decay $t \to c\gamma$ could be relevant only in a scenario with two Higgs doublets and the mass of the fourth-generation b' quark larger than m_t [45]. A most recent update addressing models I and II has found that branching ratios larger than 10^{-6} exist only for $t \to cg$ in the 2HDM-II, with $m_H = 120\,\mathrm{GeV}/c^2$ and $\tan\beta \sim 100$.

In model III [46], the above discrete symmetry is dropped and tree-level FCNC are allowed, included a tHc as displayed in Fig. 1.4(right). The neutral Higgs boson that are exchanged correspond in the MSSM model (that is a type-II model) to the h^0, H^0 and A^0. In this specific case, the tree-level FCNC decay $t \to cH$ can occur with branching ratios up to 10^{-2}. In Ref. [47], it has been found that the $t \to cH \to cWW$ rate is enhanced by several orders of magnitude with respect to its SM value, namely $BR(t \to cWW/cZZ) \sim 10^{-4}$ for an on-shell decay with $2M_W < m_H < m_t$. The same process was considered in a wider range of models, where the decay can occur through not only a scalar exchange but also a fermion or vector exchange [48]. For suitable value of heavy fermion or scalar or vectorial non-standard particle exchange, a rate as big as $BR(t \to cW^+W^-) \sim 10^{-3}$ can be reached.

General two-Higgs-doublet models are included also in Ref. [49] and an enhancement

as much as 3-4 orders of magnitude in $t \to cV$ (e.g. $BR(t \to cg) \sim 10^{-7} \div 10^{-6}$), for various values of the parameter, was declared. In the specific Type III model, values up to $BR(t \to cg) \simeq 10^{-5}$, $BR(t \to c\gamma) \simeq 10^{-7}$, $BR(t \to cZ) \simeq 10^{-6}$ are predicted [50]. Working in the most general CP-conserving 2HDM type III, with vacuum expectation values for both multiplets different from zero, a study [51] showed that driving the $\tan\beta$ properly (between 8 and 15) it is possible to enhance the branching ratio for $t \to c\gamma$ in the $(1 \div 9) \cdot 10^{-5}$ region. Similar enhancements could also appear in other rare top quark decays.

Studies on top three body-decays have been expanded by including not only threshold effects but also non-zero widths of heavy particles [52]. Working in the Type III version, branching ratio as $BR(t \to cZZ) \sim 10^{-3}$ have found.

By further extending the 2HDM sector and including a third Higgs doublets, one can give rise to a vertex HWZ at tree level in a consistent way [53]. Accordingly, the $t \to bWZ$ decay can be mediated by a charged Higgs (coupled with m_t) that can enhance the corresponding branching fraction up to a spectacular 10^{-2}.

$t \to Vq$ in SUSY models

Supersymmetry could affect the top decays in different ways. Most models consider the MSSM framework, with or without R-parity $(-1)^{3B+L+2S}$ conservation, where B and L are the leptonic and baryonic quantum numbers, and S the spin of each particle.

When the R-parity conservation is assumed, there are two sources of flavour violation in the MSSM. The first one arises from the flavour mixings of up-squarks and down-squarks, which are described by the matrix that is inherited from the Standard Model CKM. Hence charged 'electroweak-like' couplings, as those in which a top couples in a down-squarks and a chargino (depicted in Fig. 1.5, a) and b)) are possible and their amplitude adds to the SUSY Higgs boson loop-diagrams. The couplings may be larger than the W-boson loop, since the mass splitting between the squarks in the loops may be significant and the Higgsino-component couplings are non-universal Yukawa couplings. The second one results from the misalignment between the rotations that diagonalize the quark and squark sectors that, due to the presence of soft SUSY breaking terms, are not the same for the three families. This implies that, differently from the SM, neutral 'electroweak-like' couplings are possible (c) and d) in Fig. 1.5), where the top vertex is with a up-squark and a neutralino. When the loop after a neutral coupling involves a gluino, the gluon emission is possible as displayed in e) and f) in Fig. 1.5 ('QCD-like' couplings). Since such stop-scharm flavor mixings may be significant, this kind of loops involving the strong coupling may be quite sizable or dominant over other kinds of loops.

In the MSSM with universal soft breaking, generally not very large enhancements have been found. The first studies [54] considered one-loop QCD-like and charged EW-like contributions, which were later generalized in order to include the left-handed and right-handed squarks mixings [55]. These new contributions were found to enhance the Standard Model

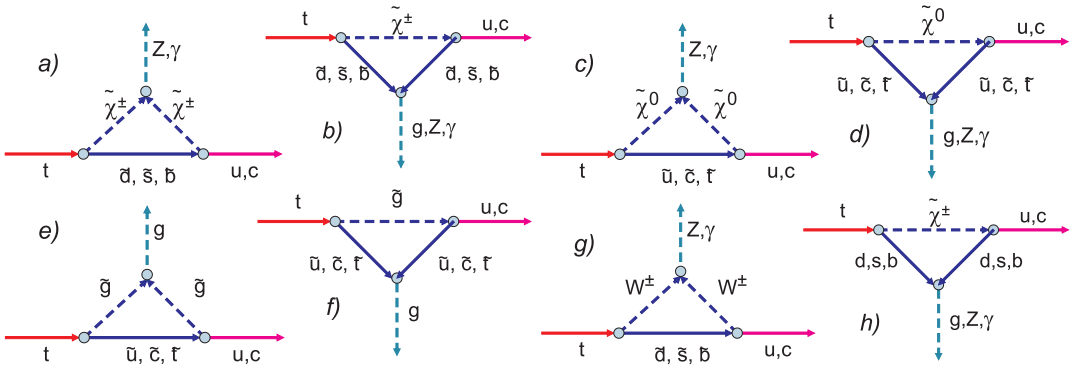

Figure 1.5: Examples of the different loop-diagrams that are possible in SUSY: charged electroweak-like (a) and b)), neutral electroweak-like (c) and d)), QCD-like (e) and f)), and R-parity violating modes (g) and h)).

branching fractions by as much as 3-4 orders of magnitude. For large values of the soft SUSY-breaking scalar mass the cZ decay mode is maximal, while a $c\gamma$ decay suppression may occur for certain combinations of parameters.

When the electroweak-like one-loop with charginos, neutralinos down-like and up-like squarks were fully included [56], it was found that such values can be reached: $BR_{MSSM}(t \to cg) \sim 10^{-5}$, $BR_{MSSM}(t \to cZ) \sim 10^{-6}$ and $BR_{MSSM}(t \to c\gamma) \sim 10^{-6}$. Few years later it was confirmed [57] that the possibility of sizable rates does not necessarily require a general pattern of gluino-mediated FCNC interactions affecting both the LH and the RH sfermion sectors, but the LH is sufficient.

The FCNC decay channel with the lightest supersimmetric Higgs ($t \to ch^0$) still seemed more promising (see for example [58]), not only because in some models the vertex tHc can be generated at tree level, but also because the GIM suppression does not apply in some loops. The BR for this decay varies between 10^{-4} and 10^{-5}, strongly depending from the $\tilde{c} - \tilde{t}$ mixing and R conservation.

At the same time, in Ref. [59] a non-universal and flavour dependent SUSY breaking was proposed, thus introducing a flavour mixing in the theory. Since there are not strong constraints on off-diagonal squark mass between the second and third families, one could envisage a situation in which there is a large mixing angle between $\tilde{c}$ and $\tilde{t}$ squarks. The amplitude of loops involving $\tilde{c}V\tilde{t}$ couplings becomes larger and rates as 10^{-5} for $t \to cV$ are obtained, that may be supposed to be detectable in next machines. The most recent study on the topic is in Ref. [60], where all top-quark FCNC processes induced by stop-scharm mixings via gluino-squark loops are fully revised. The maximal predictions have been found with a non-zero δ_{LR} parameter for the flavor mixings between left-handed scharm and stop, along with the following constraints on the SUSY masses:

$$m_{\tilde{u}} > 96 \text{ GeV}/c^2 \text{ , } m_{\tilde{u}} > 89 \text{ GeV}/c^2 \text{ , } m_{\tilde{\chi}^0} > 46 \text{ GeV}/c^2 \text{ , } m_{\tilde{\chi}^+} > 94 \text{ GeV}/c^2 \text{ ,}$$

and they read $BR_{MSSM}(t \to cg) = 1.3 \cdot 10^{-4}$, $BR_{MSSM}(t \to cZ) = 1.2 \cdot 10^{-5}$ and $BR_{MSSM}(t \to c\gamma) = 1.3 \cdot 10^{-6}$.

The models documented up to here are examples of constrained MSSM.

In Ref. [61], the branching ratios for $t \to cg/\gamma/Z$ were recalculated in an unconstrained MSSM framework. The assumptions on the soft breaking terms are relaxed and new sources of flavor violation appear in the sfermions mass matrices. In this case the neutralino-$q\tilde{q}$ and gluino-$g\tilde{g}$ couplings induce larger FCNC contributions. For favourable parameter values allowed by current precise experiments, they found $BR_{UC}(t \to cg) \sim 10^{-4}$, $BR_{UC}(t \to cZ) \sim 10^{-6}$ and $BR_{UC}(t \to c\gamma) \sim 10^{-6}$.

If the bound of R-parity conservation is removed (as in models that assume B-violating couplings), some additional loops are possible, two of which are exemplified in g) and h) in Fig. 1.5. Gauginos are not forced anymore to originate from top together with a s-quark, but couplings as $tV\tilde{q}$ and $t\tilde{V}q$ are possible. The introduction of broken R-parity models could give large enhancements [62], and make some of these channels observable. These enhancements were calculated as $BR_{\not R}(t \to cg) \sim 10^{-3}$, $BR_{\not R}(t \to cZ) \sim 10^{-4}$ and $BR_{\not R}(t \to c\gamma) \sim 10^{-5}$.

$t \to Vq$ in Technicolor

Technicolor is a typical idea to dynamically break the electroweak symmetry: here the EWSB mechanism arises from a new, strongly coupled gauge interaction at TeV energy scales. The original simple technicolor theory encounters enormous difficulty in generating fermion masses (especially the heavy top quark mass) and faced the problem of passing through the precision electroweak test. Therefore, the so-called 'top-color' scenario was proposed in order to make the predictions consistent with the LEP data and to explain the large top quark mass. Loop processes are calculated in a Farhi-Susskind one-family model [64] and in a top-color-assisted multi-scale technicolor model [65], where SM branching fractions for $t \to cV$ are enhanced by no more that 3-4 orders of magnitude.

In the more recent developments of these models (TC2, [66]), it has found that the contributions of top-pions and top-Higgs-bosons predicted can enhance the SM branching ratios by as much as 6-9 orders of magnitude. In TC2 model, FCNC interactions may be greatly enhanced if one of the following scenario verifies:

- top-color is non-universal, only causing the top quark to condensate to its large mass. The neutral top-pion has large Yukawa couplings to only top quark;
- quark masses are acquired via an 'extended technicolor' (ETC) interaction, with the exception of the top quark whose only a small portion of mass is from ETC. ETC-pions have small Yukawa couplings to all quarks, and for top quark the coupling is much weaker than the top-pions;
- if the up-quark mass matrix is composed of both ETC and top-color contributions, the simultaneous diagonalization of both the top-pions Yuwawa couplings in top-color

sector and the ETC-pions Yukawa couplings in ETC sector is not possible. Thus, after the diagonalization of the mass matrix of up-type quarks, the top-pion in top-color sector should have tree-level FCNC Yukawa couplings for the top quark.

In the extreme case, the orders of magnitude of branching ratios are $BR_{TC2}(t \to cg) \sim 10^{-3}$, $BR_{TC2}(t \to cZ) \sim 10^{-4}$ and $BR_{TC2}(t \to c\gamma) \sim 10^{-5}$ [67]. It is worth to mention that in the same TC2 model an enhancement of the $t \to cWW$ mode up to 10^{-3} is expected.

$t \to Vq$ in other SM extensions

Flavour changing neutral currents effects can be naturally induced in theories with an extended scalar sector, as well as a larger gauge group.

Alternatively to the scenario with extended Higgs sectors (in which the scalar FCNCs are induced at the tree level by the presence of additional multiplets), heavy particles lying beyond the Fermi scale can induced virtual effects leading to FCNC. A possible approach is to assume an effective Lagrangian composed of only one Higgs doublet, then extend the Yukawa sector with dimension-six $SU_L(2) \times U_Y(1)$-invariant operators. The model produces CP-even and CP-odd tVq vertexes at the tree level, with branching ratio for the Higgs-mediated decays enhanced by two or three orders of magnitude with respect to the 2HDM results. As discussed in Ref. [63], the rates for $t \to cV$ are proportional to $1/\Lambda^2$ (with Λ new physics scale) and vary smoothly with the Higgs boson mass in the range $120 < m_H < 170\,\mathrm{GeV}/c^2$. The maximal values in this effective Lagrangian approach (ELA) are $BR_{ELA}(t \to cg) = 3.4 \cdot 10^{-6}$, $BR_{ELA}(t \to cZ) = 2.4 \cdot 10^{-5}$ and $BR_{ELA}(t \to c\gamma) = 1.7 \cdot 10^{-7}$, obtained when $\Lambda = 400\,\mathrm{GeV}$.

Left-Right (LR) symmetry models are based on the gauge group $SU_L(2) \times SU_R(2) \times U_{B-L}(1)$ and aimed to understand the origin of parity violation in low energy weak interactions. An extension including vector-like heavy fermions induces non-unitarity of the CKM matrix, thus FCNCs are possible at tree-level. Favorable enhancements were found in an unconstrained LR SUSY model, where breaking parameters are allowed to induce flavor-dependent mixings in the squark mass matrix, so they can be arbitrarily large between the second and third generations: $BR_{LR}(t \to cg) \sim 10^{-4}$, $BR_{LR}(t \to cZ) \sim 10^{-5}$ and $BR_{LR}(t \to c\gamma) \sim 10^{-6}$ [68].

CKM matrix unitarity is loosed also in models with extra quarks and tZc, tHc may arise at the tree level. If new quarks are $SU_L(2)$ singlets with $Q = 2/3$, present experimental data allow $BR_{QS}(t \to cZ) \sim 1.1 \cdot 10^{-4}$ while $BR_{QS}(t \to cg) \sim 1.5 \cdot 10^{-7}$, $BR_{QS}(t \to c\gamma) \sim 7.5 \cdot 10^{-9}$ [40, 69]. Even $t \to cH$ is raised up to $\sim 4 \cdot 10^{-5}$ level. In models with $Q = -1/3$ quark singlets, the respective branching ratios are much smaller since the breaking of the CKM unitarity is very constrained by experimental data.

Recent theoretical elaborations proposed models where the decay rate into a Higgs boson is enhanced. In Ref. [70] an effective 4D theory is discussed, containing fundamental SM fermion and gauge fields with composite string-color fields of extra dimensional origin. One

of the by-product of this scenario is a branching ratio for $t \to ch$ (with h the lightest Higgs) as large as 10^{-4}.

As a final remark, it is worth to be noticed that the appearance of flavour changing neutral currents is often an undesired effect in models beyond the SM at the TeV scale, because it does not meet the present experimental evidences. The most popular approach is to invoke the principle of Minimal Flavour Violation (see for instance Ref. [71]), that can be used in theories with low energy supersymmetry, multi-Higgs doublet and others and can be implemented in Grand Unified Theories.

Summary of $t \to Vq$ predictions and perspectives

In scenarios beyond the SM, enhancements in FCNC decays arise either from a large virtual mass or from the couplings involved in the loop. A summary for the main models outlined above (2HDM in the type-II and III, SUSY with and without R-parity conservation, top-assisted technicolor, left-right asymmetry model and quark singlets) is in Tab. 1.5. The most up-to-date predictions presented in the text are quoted, taking the maximal values separately for each decay in each model. When the limit needs a specific parameter choice different from other channels, it is explicitly noticed. Only the order of magnitude is quoted here.

Table 1.5: Branching ratios for FCNC top quark decays as predicted within the in seven SM extensions. Details on the specific models and more precise upper limits are in the text.

	2HDM-II	2HDM-III	MSSM with R	MSSM with $\not{R}$	TC2	LR	QS
$BR(t \to qg)$	10^{-5}	10^{-4}	$10^{-4\,(*)}$	10^{-3}	10^{-3}	10^{-5}	10^{-7}
$BR(t \to q\gamma)$	10^{-7}	$9 \cdot 10^{-4\,(**)}$	10^{-6}	10^{-5}	10^{-7}	10^{-6}	10^{-8}
$BR(t \to qZ)$	10^{-8}	10^{-6}	$10^{-5\,(***)}$	10^{-4}	10^{-4}	10^{-4}	10^{-4}

(*): in an unconstrained MSSM model.
(**): with $8 < \tan\beta < 15$.
(***): with constraints in Ref. [60].

The analyses presented in next chapters will show that branching ratios down to 10^{-4} are reachable from CMS, improving to some 10^{-5} after integrated luminosity of hundreds inverse femtobarns and a combination with ATLAS results for the same channels. Such sensitivities will be reached only for the $t \to qZ$ and $t \to q\gamma$ decay modes.

Multi-Higgs doublets models and MSSM generally are not the ones producing the largest FCNC rates, but some specific configurations could lead to huge enhancements. It is widely accepted that hints of supersymmetry could be likely to appear in the very initial phase of LHC physics run, via the detection of inclusive multi-lepton, multi-jet and missing energy signature. Searches for several MSSM Higgs decay modes are among the main target of both

CMS and ATLAS experiments. On the other hand, there are no clean ways to estimate the values of $\tan\beta$, $\tilde{c}$-$\tilde{t}$ mass splitting and other supersymmetry parameters, at least in the low luminosity phase. SUSY-induced FCNC couplings have instead a significant sensitivity to such parameters, thus putting some limits on neutral top decays could do a better job in constrain that parameters. In Ref. [60], for instance, it is shown how branching ratios for $t \to cZ/\gamma$ drop as sparticles become heavy, enhance with the quantity $X_t = A_t - \mu \cot\beta$ (that affects the squark mass splittings) and is very sensitive to mixing parameters as δ_{LL} and δ_{LR}, that is a very peculiar feature of these modes.

Top flavour changing decays could be also a way to constrain the amount of R-parity non conservation. While in direct searches of SUSY partners the study of this conservation would require an exclusive reconstruction of both the squarks and gauginos final state, Tab. 1.5 shows that the $t \to qZ$ branching ratio could provide a much simpler tool. Since values as high as 10^{-4} are expected only for R-parity broken modes, reaching a better exclusion limit would rule out the violation. Further interesting possibilities could be offered by pushing the sensitivity below some 10^{-4}, because it would shed light on more exotic models especially in the qZ sector. While the production of top-color states seems unlikely to be reached, the expectations from the quark-singlet models may offer a more viable mean to tag the existence of a charge-2/3 heavy T quark, for instance with respect to the $T \to tZ$ direct search [72]. Finally, left-right models predict the existence of a Z' heavy boson (Z_{LRM} or Z_{ALRM}) in the TeV region, that would produce a striking signal event at the very initial operation phase, than would be precisely constrained after the first inverse femtobarns. The observation of this kind of objects would further motivate the search of a top FCNC decay at the 10^{-4} branching ratio level.

As for the decays with the c quark replaced by the u one, the respective branching ratios are smaller by a factor of m_u/m_c, and thus they would be out of the LHC reach.

1.3.3 Current experimental limits on tVq couplings

The purpose of this section is to report about the current experimental limits on the tVq couplings. They can be derived both from present (Tevatron) and recent past (HERA and LEP 2) machines, and from constraints deduced from observables in other sectors – as electroweak and CP-mixing parameters.

Anomalous tVq couplings could well enter the top quark production vertexes too and they are expected to be enhances in beyond SM scenarios [73]. This interference may lead to a modification of the production (singly or pairly) cross sections, or to some characteristic top production signatures. Some of the possible Feynman diagrams for neutral vertexes in top production are depicted in Fig. 1.6.

The most studied channel in all collider experiments is the one leading to a top-charm associated production, because it can be effectively identified in e^+e^- and ep collisions.

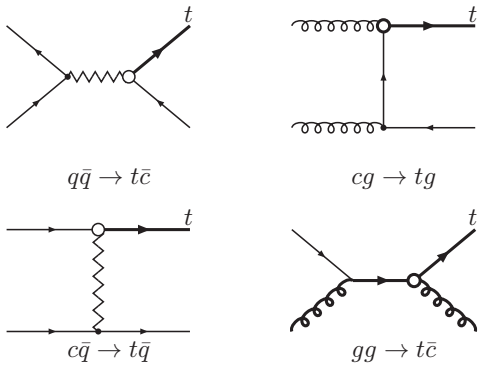

Figure 1.6: Anomalous processes occurring in the single-top production. Circles indicates the tVq vertexes. In the first and third diagrams, vector boson can be a g, Z^0 or γ.

Constraints from EW and CP-asymmetry observables

The measurement of the inclusive branching ratio for the FCNC process $b \to s\gamma$ has been used to put constraints on the $t\gamma c$ and tgc couplings [74]. The known branching ratio for $t \to bW$ and the $b \to s\gamma$ bounds from the CLEO experiment place the limits $\kappa_{tc}^{g} < 0.9$ and $\kappa_{tc}^{\gamma} < 0.16$, which can be translated into the upper limits $BR(t \to cg) < 3.4 \cdot 10^{-2}$ and $BR(t \to c\gamma) < 2.2 \cdot 10^{-3}$.

The tZc couplings was bounded using several FCNC low-energy processes such as $K_L \to \mu^+\mu^-$, $K_L - K_S$ mass difference, $B_0/\bar{B}_0$ mixing and $B \to l^+l^-\gamma$ (that is linear in tZc coupling), as well as the oblique parameters ρ and S. Combining these bounds with precision EW observables as Γ_Z, R_c, R_l, R_b one obtains $BR(t \to cZ) < 1.6 \cdot 10^{-2}$ at 95% C.L, when $116 < m_H < 170\,\mathrm{GeV}/c^2$ [75]. On the other hand, the contribution of EW observables on $t \to c\gamma$ are scaled by a factor $(M_Z/m_t)^4$, thus this channel does not receive significant constraints from them.

Experimental results from Tevatron

On the experimental side, the Tevatron $\bar{p}p$ collider (having a $\sqrt{s} = 1.8 \div 1.96\,\mathrm{TeV}$) offers one of the greatest opportunity to search for anomalous couplings, both in top decay and production. This machine is currently in a steady operation phase, thus the current results reported here may well be updated by further analyses, following the day-by-day increase of integrated luminosity.

CDF and DØ experiments have addressed the search of anomalous top couplings both in the $t \to qg/\gamma/Z$ decays (mostly in the $t\bar{t}$ production) and in single top production, leading to a $t\bar{q}$ in the final state.

For the $t \to \gamma c(u)$ and $t \to Zc(u)$ decays, the following 95% C.L. upper limits published by CDF [13] are:

$$BR(t \to c\gamma) + BR(t \to u\gamma) < 3.2\%,$$
$$BR(t \to cZ) + BR(t \to uZ) < 33\%,$$

that have to be intended as $BR(t \to qZ/\gamma) = \Gamma(Z/\gamma q)/\Gamma(Wb)$, with anomalous couplings occurring one at the time. Recently, a statistic as large as 1.12 fb^{-1} has been integrated and this upper limit has been updated to 10.6% [76].

These results can be translated into bounds on the top anomalous couplings:

$$\kappa^{\gamma}_{tq} < 0.76,$$
$$\sqrt{|v^{Z}_{tq}|^2 + |a^{Z}_{tq}|^2} < 0.59, \tag{1.8}$$

where, as usually done, $\Lambda = 1\,\text{TeV}$. If the integrated luminosity would be extrapolated up to the higher values expected from Tevatron (5-6 fb^{-1} by the end of 2009), these limits would extend to $BR(t \to qZ) < 4.6\%$ and $BR(t \to q\gamma) < 0.44\%$.

The search for processes such as those displayed in Fig. 1.6 exploits the leptonic W decay from the top, along with an anti-b-tagged jet well-separated from the lepton. CDF results obtained with 1 fb^{-1}, give coupling limits of $\kappa^{g}_{tu} < 0.058$ and $\kappa^{g}_{tc} < 0.22$ at 95% C.L.. A moderate improvement on tgq with respect to CDF limits is obtained [77] from DØ by analyzing 230 pb^{-1} of lepton+jets inclusive data: $\kappa^{g}_{tu} < 0.037$ and $\kappa^{g}_{tc} < 0.15$ at 95% C.L..

Further constraints on the tgq vertex was derived from the study of the $t\bar{t}$-pair production cross section, comparing the measured value $\sigma(t\bar{t})$=6.7 $\pm$ 1.3 pb, [2, 3] with the expected one. Imposing that the $t\bar{t}$ production cross section, including the possible effect of anomalous couplings, should not differ from the observed one by more than 2 pb, leads to the constraint [78] $\kappa^{g}_{tq} < 0.47$, that is consistent with results from single-top production.

Experimental results from LEP 2

Since 1997, LEP 2 has ran at $\sqrt{s}$ exceeding of 180 GeV, making the production of single top quark kinematically possible through the reaction:

$$e^+e^- \to \gamma^*/Z^* \to \bar{t}c(u).$$

The LEP experiments [79] have presented the results of their search for this process. Upper limits around $0.3 \div 0.6$ pb are found in a range of energy for a given integrated luminosity. Assuming $m_t = 175\,\text{GeV}/c^2$ and considering tZq and $t\gamma q$ different from zero one at the time, these cross section limits can be converted in bounds for branching ratio, then to $t\gamma/Zq$ upper limit. When data in the $\sqrt{s} = 189 \div 209$ GeV energy domain are cumulated for ALEPH (214 pb^{-1} [80]), DELPHI (541 pb^{-1} [81]), L3 (634 pb^{-1} [82]) and OPAL (600 pb^{-1} [83]), the results in Table 1.6 are obtained.

Table 1.6: Current constraints on top quark FCNC interactions from LEP 2 results.

$t \to \gamma q$	$BR < 3.2\%$	$\kappa_{tq}^{\gamma} < 0.76$	(other FCNC couplings zero)
$t \to Zq$	$BR < 14\%$	$\sqrt{\lvert v_{tq}^{Z}\rvert^2 + \lvert a_{tq}^{Z}\rvert^2} < 0.38$	(other FCNC couplings zero)

Table 1.7: Results from the two experiments at the HERA ep collider, for the search of a FCNC single-top production. It is assumed $m_t = 175\,\mathrm{GeV}/c^2$ and $BR(t \to qZ) = 0$.

Collab.	$\sqrt{s}$ (GeV)	$\mathcal{L}(\mathrm{pb}^{-1})$	BR$(t \to \gamma u)$(95% C.L.)	κ_{tu}^{γ} (95% C.L.)
ZEUS [84]	$300 - 318\,\mathrm{GeV}$	130.1	< 0.0059	< 0.328
H1 [85]	319 GeV	118.3	< 0.0132	< 0.491

The recent CDF upper limit on tZq (to be published) has improved the LEP estimate, so it is the new candidate for the best limit for this coupling. On the other side, $t\gamma q$ results much better constrained by HERA limits, as mentioned below.

Experimental results from HERA

In the HERA ep collider ($\sqrt{s} = 318\,\mathrm{GeV}$) top quark can be singly produced either by a charged current interaction ($ep \to \nu t\bar{b}X$) or a FCNC process $e^{\pm}p \to e^{\pm}p(t/\bar{t}) + X$, which includes the FCN interactions $Z^*/\gamma^* \to t\bar{q}$ (the Z^0 exchange is suppressed by the large propagator mass). Since the SM cross section for single-top production is less than 1 fb, any observed single-top event in the HERA data would be a clear sign of physics beyond the SM.

An inclusive search for FCNC single-top production has been performed for H1 and ZEUS experiments. While no evidence of such processes are seen by ZEUS [84], in 2004 H1 found 5 events in the e/μ channels, being 1.31 ± 0.22 the number expected if only the Standard Model would be at work [85]. No excess was found for the hadronic channel. The observed cross section is $\sigma(ep \to etX) = 0.29^{+0.15}_{-0.14}$ at $\sqrt{s}$ =319 GeV. If intended as a statistical fluctuation, this value corresponds to 0.55 pb 95% upper limit. This result can be converted in a limit for the $t\gamma u$ anomalous coupling and in a branching ratio for $t \to \gamma q$, by assuming the $BR(t \to qZ) = 0$.

A summary of these data is given in Tab. 1.7 and can be compared with Tevatron and LEP 2 results. Since the u-quark density of the proton is much higher than the c-quark density, the production of single top quarks is most sensitive to a coupling of the type $t\gamma u$.

Due to the excess reported from H1, the result from ZEUS is better and currently the best limit for the $t\gamma u$ coupling.

The whole experimental situation is depicted in Fig. 1.7, where the exclusion limits at 95% for the decay channels addressed in the work are indicated.

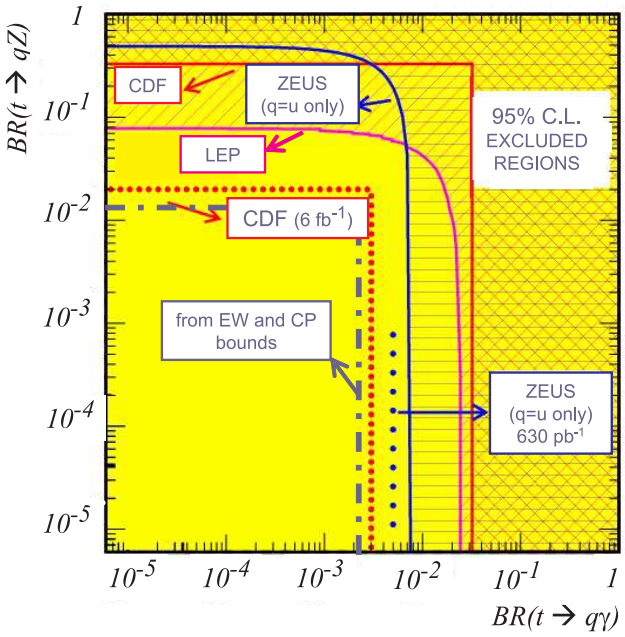

Figure 1.7: The 95% C.L. exclusion plot for the current (and recent past) experimental limits to $t \to q\gamma$ and $t \to qZ$ branching ratios. It can be considered as a picture of the scenario before the venue of LHC experiments.

During the 90s the first limits on $t \to qZ/\gamma$ were put by the workhorses LEP 2 (pink line) and Tevatron Run-I (red line), reaching quite similar results for a direct search of the $t \to q\gamma$ decay. Today, the huge number of top pairs integrated at Fermilab allows to reach the LEP 2 sensitivity even in the $t \to qZ$ mode. None of these machines have claimed the observation of any anomalous decay. If the results from the (recently dismantled) HERA collider are put in the game and the excess declared from H1 is intended as a statistical fluctuation, limits on $t\gamma u$ are significantly enhanced (blue line). It is worth to note that the bounds determined from CP and EW observables (blue dashed-dotted line) have still to be reached, and they will not be even at the end of the Tevatron life (red dotted line). In other words, if an evidence of a $t \to qZ/\gamma$ would be declared today, a full revision of measure in the LEP observables and in the K and B CP-mixings would be required, along with the underlining theories. On the contrary, pushing the experimental sensitivity under these limits will imply that a top FCNC decay can occur without corrections on parameters in other sectors. LHC experiments will demonstrate to have such a good sensitivity.

The tgq coupling is not addressed in the present work, mainly because the $t \to gq$ is not among the most promising channels. A dense multi-jet environment will be the most striking feature of the LHC collisions, thus few chances will exist to efficiently identify the jet coming from the gluon from a top decay. Different phenomenologies have been proposed to tag this anomalous vertex, as the production of like-sign top pairs tt or $\overline{tt}$ [22, 86]. This is indeed a

'spectacular' and quite unambiguous signal, almost free of background.

1.3.4 Previous estimates of FCNC observability at LHC

The LHC collider is presented in the next chapter, were a quick overview of the CMS experiment will be provided. Since the first studies on the ATLAS and CMS physics capabilities, there has been clear evidence that branching ratios for the top decay as low as $10^{-4} \div 10^{-5}$ can be reached. As explained above, there are some beyond SM scenarios leading to such rates, hence the possibility to detect a top FCNC channel was addressed by various groups to some extent.

The first realistic studies were performed used the so-called "fast simulations", that attempt to simulate the detector response using parametrized resolution functions. Some insights of the CMS simulation package will be provided in the next chapter.

ATLAS was the first collaboration to present its Physical Technical Design Report [87], where most of the work was done using the fast simulation framework ATLFAST [88]. The sensitivity to the decay $t \to Zu/c$ has been analyzed (see also Ref. [89]) by searching for a signal in the channel $t\bar{t} \to (Zq)(W\bar{b})$, with the boson being reconstructed via the leptonic decay $Z \to ll$. Both hadronic and leptonic W decay modes were considered. The signal was generated by artificially changing the $t \to sW$ in the $t \to cW$ mode and the only backgrounds considered was Z+jets, WZ and $t\bar{t}$ production (only in the leptonic mode), generated via PYTHIA 5.7 [90]. The selection cuts required a pair of hard isolated, opposite sign, same flavor leptons (electrons or muons), forming a Z in an 6 GeV/c^2 wide invariant mass window. With a proper selection of the b-jets and missing transverse energy, along with constraints on the W and $M(bW)$ invariant mass, a 5σ limit of $2.0 \cdot 10^{-4}$ and $5.9 \cdot 10^{-4}$ for the leptonic and hadronic modes respectively was found, with a simulated statistic corresponding to 100 fb^{-1}.

The sensitivity to the decay $t \to \gamma u/c$ was studied by searching for a peak above background in the $M(\gamma j)$ spectrum in the region of m_t. The requirement of a high p_T isolated photon candidate, along with a $t \to Wb \to l\nu b$ reconstructed process was imposed, adopting the same simulation and reconstruction tools of $t \to Zq$. The background process considered were the pair and single-top production, $W + b\bar{b}$ and W+jets. The reconstruction tools were modified in order to produce a more conservative jet rejection, and the result was $BR(t \to q\gamma) < 1.0 \cdot 10^{-4}$ at 5σ after 100 fb^{-1}.

The CMS studies addressing the same channels followed and they exploited a dedicated Monte Carlo generator called TOPREX [91], that is able to account for the spin correlations in the objects produced from top. The detector was simulated with a fast simulation package (CMSJET, [92]) and the b-tagging efficiency was fixed at 60% (with a c and $udsg$ mistagging at 10% and $1 \div 2$% respectively). The set of background sources included single and pair top production, di-boson and Z+jets and the set of cuts was similar to ATLAS.

The declared upper limits (for a 5σ discovery level and 100 fb^{-1} of integrated luminosity) were quite similar for the Z channel ($BR(t \to u/cZ) < 1.9 \cdot 10^{-4}$) and much better for the

photonic one ($BR(t \to u/c\gamma) < 3.4 \cdot 10^{-5}$) [22, 93].

In the meantime, CMS collaboration has put many efforts in regularly updating its software framework, in order to produce the most realistic scenario for the detector simulation and the physics objects reconstruction. The full simulation environment (ORCA, [94]) was extensively adopted to produce all the results in the CMS Physical Design Report [95]. In this framework, a detailed study was performed by the author of the present work *et al.* [96]. A larger set of background is considered (as the important $Z + b\bar{b}$ and the multi-jet production) and most important systematic effects were included. This leads to the reachable upper limits (at 5σ discovery level) $BR(t \to qZ) < 14.9 \cdot 10^{-4}$ and $BR(t \to q\gamma) < 8.4 \cdot 10^{-4}$ after the first $10\,\mathrm{fb}^{-1}$, that can be extrapolated to $BR(t \to qZ) < 3.1 \cdot 10^{-4}$ and $BR(t \to q\gamma) < 2.5 \cdot 10^{-4}$ at $100\,\mathrm{fb}^{-1}$. The work presented here is intended as a large extension of that analysis, that is updated and fully motivated.

At the same time, ATLAS collaboration has revised its strategy for signal extraction by proposing a probabilistic approach [97]. For the $t \to qZ$ signal, preselected events with a reconstructed Z, large E_T^{miss} and the two hardest jets (one b-tagged) are used to build a discriminant variable (likelihood ratio) $L_R = \Pi_i P_i^S / \Pi_i P_i^B$. P_i are the p.d.f. of the following quantities: the $M(l^+l^-)$ with the largest p_T out of the 3 leptons, p_T of lepton not from Z, $M(jl^+l^-)$ and p_T of the leading b-jet. For the $t \to \gamma q$ case, preselected events are required to have one b-tag among two hardest jet and the leading γ with $p_T(\gamma) > 75\,\mathrm{GeV}/c$. A similar L_R is built with p.d.f. for $p_T(\gamma)$, $M(j\gamma)$ and number of jets. From these shapes of L_R one can argue that likelihood ratios can be used as discriminant variable. An upper cut on both L_R's is performed, choosing the point where $S/\sqrt{B}$ is the largest. Though this method is significantly different from the cut-based analysis adopted by CMS, the result are fairly similar for the Z channel, reading $BR(t \to qZ) < 13.0 \cdot 10^{-4}$ after the first $10\,\mathrm{fb}^{-1}$ and $BR(t \to qZ) < 4.1 \cdot 10^{-4}$ at $100\,\mathrm{fb}^{-1}$. The results for the photonic channel are $BR(t \to q\gamma) < 1.6 \cdot 10^{-4}$ after $10\,\mathrm{fb}^{-1}$ and $BR(t \to q\gamma) < 5.1 \cdot 10^{-5}$ at $100\,\mathrm{fb}^{-1}$, and the difference arises in part because systematic effects are taken into account differently.

It is important to observe that in these latter studies the pile-up for the low-luminosity phase is considered, thus extending the result to the $100\,\mathrm{fb}^{-1}$ scenario may not be so straightforward. In the work presented here, such extrapolation will be carefully discussed.

Finally, the possibilities for a FCNC detection in top decays has been included in a study of the physical potential for Super LHC [98]. This LHC upgrade, that has an enhanced luminosity and will be briefly presented in the Appendix (Sec. A.3), should improve the sensitivity by one or two orders of magnitude. The cuts proposed for this analysis were developed in parallel to the first CMS fast simulation and closely resembles them. Since the ability to identify the b-jets with a secondary vertex technique could be an issue in the highest luminosity environment, different tagging efficiencies were considered. With the standard 60% efficiency for b-tagging and 10%/2% for c/dsg-mistagging, results were $BR(t \to q\gamma) < 0.88 \cdot 10^{-5}$ after $600\,\mathrm{fb}^{-1}$ and $BR(t \to q\gamma) < 0.26 \cdot 10^{-5}$ at $6000\,\mathrm{fb}^{-1}$, and $BR(t \to qZ) < 1.1 \cdot 10^{-5}$

after $600\,\mathrm{fb}^{-1}$ and $BR(t \to qZ) < 0.11 \cdot 10^{-5}$ at $6000\,\mathrm{fb}^{-1}$. On the other hand, if tagging capability would be badly worsened, only techniques based on semileptonic muon decays of b-quark could be adopted. In such a case, sensitivities downgrade by a factor ≈ 4 for the $t \to q\gamma$ case and ≈ 75 for the $t \to qZ$. The present work will show that, in such a context, and intermediate situation could present.

Perspectives for the next Linear Collider, that appears to operate in a rather far future and presents a quite different environment, are not addressed here.

Chapter 2

The CMS detector

The Large Hadron Collider (LHC) [99], at the CERN Laboratory (the European Laboratory for Particle Physics, outside Geneva, Switzerland) will be completed by the early 2008. The LHC will be the highest energy accelerator in the world for many years following its completion, thus a unique tool for fundamental physics research.

Although the Standard Model of particle physics has so far been tested to exquisite precision, it is considered to be an effective theory up to roughly one TeV. The prime motivation of LHC is to elucidate the nature of Electroweak Symmetry Breaking (EWSB), for which the Higgs mechanism is presumed to be responsible. The experimental study of the Higgs mechanism can shed light on the mathematical consistency of the SM at energy scales above about 1 TeV. A wide range of alternatives that invoke more symmetry (such as supersymmetry) or new forces or constituents (such as strongly-broken electroweak symmetry or technicolor) will be extensively addressed. Undoubtedly, even some unknown mechanism could manifest at the TeV energy scale, thus all efforts to explore this scale seem well motivated.

This chapter is intended to introduce the experimental environment in which analysis is performed as well as the adopted tools. Section 2.1 describes the general features of the accelerator, LHC, and present the concept of CMS, remarking the challenges in which it is embarked. Section 2.2 sketches the concept design of CMS apparatus, its main sub-detectors and functionality of the trigger system. The interested reader can find some more details on different sub-detectors, requirements and performance in Sec. A.2.7. Subsequent section gets closer to the analysis level: it contains some details about how the objects relevant for the present work (electrons and photons, muons and missing energy, jets, jets with beauty flavour) can be reconstructed in CMS. Both the on-line algorithms (running in the trigger) and the off-line strategy (adopted by the reconstruction tools) are presented. These specific reconstruction tools, along with the programs employed for the CMS simulation and digitization chain, are mentioned in the last section.

2.1 The LHC machine and CMS experiments

Hadron colliders are well suited tools for exploration of new energy domains. If the beams energy and the luminosity are high enough, proton constituents can collide with a center-of-mass energy about 1 TeV. The LHC will provide two proton beams, circulating in opposite directions, at an energy of 7 TeV each (center-of-mass $\sqrt{s} = 14$ TeV) with a design luminosity ranging from $2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$ (low luminosity phase) to $10^{34}\text{cm}^{-2}\text{s}^{-1}$ (high luminosity phase). This corresponds to a seven-fold increase in energy and a hundred-fold increase in instantaneous luminosity over the current hadron collider experiments.

2.1.1 The LHC relevant numbers

The machine parameters relevant for the operation of CMS are listed in Tab. 2.1.

Table 2.1: The machine parameters relevant for the LHC, regarding the beam and the collisions. Only the operation with p-p is concerned here.

Energy per nucleon	E	=	7 TeV
Dipole field at 7 TeV	B	=	8.33 T
Design Luminosity	$\mathcal{L}$	=	$10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$
Bunch separation			25 ns
No. of bunches	k_B	=	2808
No. particles per bunch	N_p	=	1.15×10^{11}
β-value at IP	β^*	=	0.55 m
RMS beam radius at IP	σ^*	=	$16.7\,\mu m$
Luminosity lifetime	τ_L	=	15 hr
Number of collisions/crossing	n_c	$\approx$	20

The LHC machine comprises 1232 dipole magnets, a set of other special magnets and r.f. cavities providing a"kick" that results in an increase in the proton energy of 0.5 MeV/turn. The luminosity is given by:

$$\mathcal{L} = \frac{\gamma\, f\, k_B\, N_p^2}{4\pi\, \epsilon_n\, \beta^*} F,$$

where γ is the Lorentz factor, f the revolution frequency, k_B the number of bunches, N_p the number of protons/bunch, ϵ_n the normalized transverse emittance (with a design value of $3.75\mu m$), β^* the betatron function at the interaction point, and F the reduction factor due to the crossing angle. With a nominal energy of each proton of 7 TeV and a design luminosity of $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$, around 1 billion proton-proton interactions per second are delivered.

Table 2.2: Approximate event rates of some physics processes at the LHC for the low luminosity phase ($2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$). In this table, one year is equivalent to $20\,\mathrm{fb}^{-1}$.

Process	Events/s	Events/year
$W \to e\nu$	40	$4 \cdot 10^8$
$Z \to ee$	4	$4 \cdot 10^7$
$t\bar{t}$	1.6	$1.6 \cdot 10^7$
$b\bar{b}$	10^6	10^{13}
$\tilde{g}\tilde{g}$ (m = 1 TeV/c^2)	0.002	$2 \cdot 10^4$
Higgs (m_H = 120 GeV/c^2)	0.08	$8 \cdot 10^5$
Higgs (m_H = 800 GeV/c^2)	0.001	$8 \cdot 10^4$
QCD jets ($p_T > 200$ GeV/c)	10^2	10^9

The bunches are formed in the 26 GeV Proton Synchrotron (PS) with a 25 ns spacing. The beam is subsequently accelerated to 450 GeV in the Super Proton Synchrotron (SPS) and transferred to the LHC. When the machine will steadily operate at a bunch spacing of 25 ns, the number of bunches will be the nominal $k_B = 2808$ with a crossing angle of 285 μrad. The beam current cannot exceed half the nominal value, since part of the beam dump and collimation systems are staged. This will limit the initial luminosity to $\mathcal{L} = 2 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ until the 2010 run. This nominal luminosity is the default one for the bulk of the present analysis, and most of simulations were performed assuming this value.

An integrated luminosity of $10\,\mathrm{fb}^{-1}$, that is the baseline choice for the present work, should in principle be reached in less than one year run. In a realistic scenario, the integrated luminosity will most likely be limited by the time taken to master LHC operation and inefficiencies, thus results contained here could be obtained in a quite longer running time.

2.1.2 The CMS experiment: the physics programme

The CMS experiment [100] is one of the general purpose detector at the LHC, aimed to explore physics at an unprecedent energy scale. Beside the investigations in the Higgs sector and new physics, CMS will also be an instrument to perform precision measurements of Standard Model parameters, mainly as a result of the very high event rates. Among the measurements, careful studies on QCD, electroweak and flavour physics will extend our current knowledge and can complement direct searches with indications for some new physics.

The few processes listed in Tab. 2.2 have a very large rate even in the low luminosity phase, when they are expected to be essential. The LHC will be a Z factory, a W factory, a b quark factory, a top quark factory – and even a Higgs or SUSY s-particle factory if these new particles have TeV scale masses.

As throughly explained in the previous chapter, the present work resides on the large $t\bar{t}$

production cross section (a rate measured in Hz), and aims at the evaluation of the impact of LHC on a specific top physics sector.

The CMS experimental challenge

In order to meet the goals of the LHC physics programme, very stringent requirements are demanded to the CMS global performances. Most of them are particularly important for the reconstruction of Standard Model objects and top quark final states. They can be summarized as follows:

- good muon identification and momentum resolution, especially from 'central' muons ($|\eta|$ <2.5[1]) coming from W and Z boson (transverse momentum up to 100-200 GeV/c). The ability to determine unambiguously the charge of muons and a di-muon mass resolution around 1% at the Z scale are equally important;

- good reconstruction efficiency and momentum resolution in the inner tracker, where the transverse momentum of all the charged particle is measured;

- good electron identification up to the kinematic range of electron from W and Z, photon identification and energy measurement in a wide energy spectrum. This requires a very good electromagnetic energy resolution, the capability to reconstruct the photon direction and the primary interaction vertex, a good di-photon and di-electron mass resolution (around 1% at the Z scale) and an efficient rejection of the $\pi^0 \to \gamma\gamma$ photon background;

- as best as possible efficiency in tagging the jets coming from a b quark (hereafter "b-jets"). This results in a pixel detectors close to the interaction region;

- good missing transverse energy resolution, in order to distinguish effectively the W decays from background processes. Hadron calorimeters have to feature a hermetic geometric coverage ($|\eta|$ <5) and a fine lateral segmentation ($\Delta\eta \times \Delta\phi < 0.1 \times 0.1$[2]).

CMS has been designed to meet all these requirements. The main distinguishing features, detailed in the next section, are a high-field solenoid, a full silicon-based inner tracking system and a fully active scintillating crystal-based electromagnetic calorimeter.

General requirements to the overall apparatus

At $\sqrt{s} = 14\,\text{TeV}$ protons collide with a total cross section around 100 mb. When the luminosity will reach the design value, an event rate of the order of 10^9 inelastic events/s will be

[1] The definition of η will be given at the end of this section.
[2] The definition of ϕ will be given at the end of this section.

observed. This extraordinary particle multiplicity challenges the present detection capability, imposing some additional requirements that detectors and readout electronic cannot fail. They are:

1. the on-line event selection process ("trigger") must reduce the approximately 1 billion interactions/s to no more than about 100 events/s, manageable for storage and subsequent analysis. The short (25 ns) time between bunch crossings has major implications for the design of the readout and trigger systems;

2. at the design luminosity, a mean of about 20 inelastic collisions will be overlapped on the event of interest. This implies that around 1000 charged particles will emerge from the interaction region every 25 ns. The products of an interaction under study may be confused with those from other interactions in the same bunch crossing: when the response time of a detector element and its electronic signal is longer than 25 ns, the problem becomes severe. The effect of this "pile-up" can be reduced by using high granularity detectors with good time resolution, resulting in low percentage of readout channels fired by the same bunch- crossing (low "occupancy"). This requires a large number of detector channels (of the order of millions). A big effort in time synchronization is just one of the challenges resulting from this choice;

3. the large flux of particles coming from the interaction region leads to high radiation levels, requiring both radiation-hard detectors and front-end electronics.

The designed apparatus is conceived to have all the capabilities listed above. A sketch of the detectors and trigger system that has been designed – and today is on the way to be completed soon – is outlined in next sections and in Sec. A.2.7.

Coordinate conventions

The coordinate system adopted by CMS follows the standard for these cylinder-shaped experiments. It has the origin centered at the nominal collision (located in the geometrical center of the detector), the y-axis pointing vertically upward and the x-axis pointing radially inward toward the center of the LHC. Thus, the z-axis points along the beam direction (anticlockwise). The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis. The rest frame of the hard collision is generally boosted relative to the laboratory frame along the beam direction, so it is worth to work in a space that is invariant under boosts along this direction. Such space has designed to be the (η, ϕ) space, where the pseudorapidity η is related to the polar angle θ and defined as $\eta = -\ln(\tan(\theta/2))$.

The momentum measured in the plane transverse to the beam direction (p_T, "transverse momentum") is given by $p_T = p\sin\theta = \sqrt{p_x^2 + p_y^2}$, and similarly for the transverse energy E_T. The imbalance of energy measured in the transverse plane is denoted by E_T^{miss} .

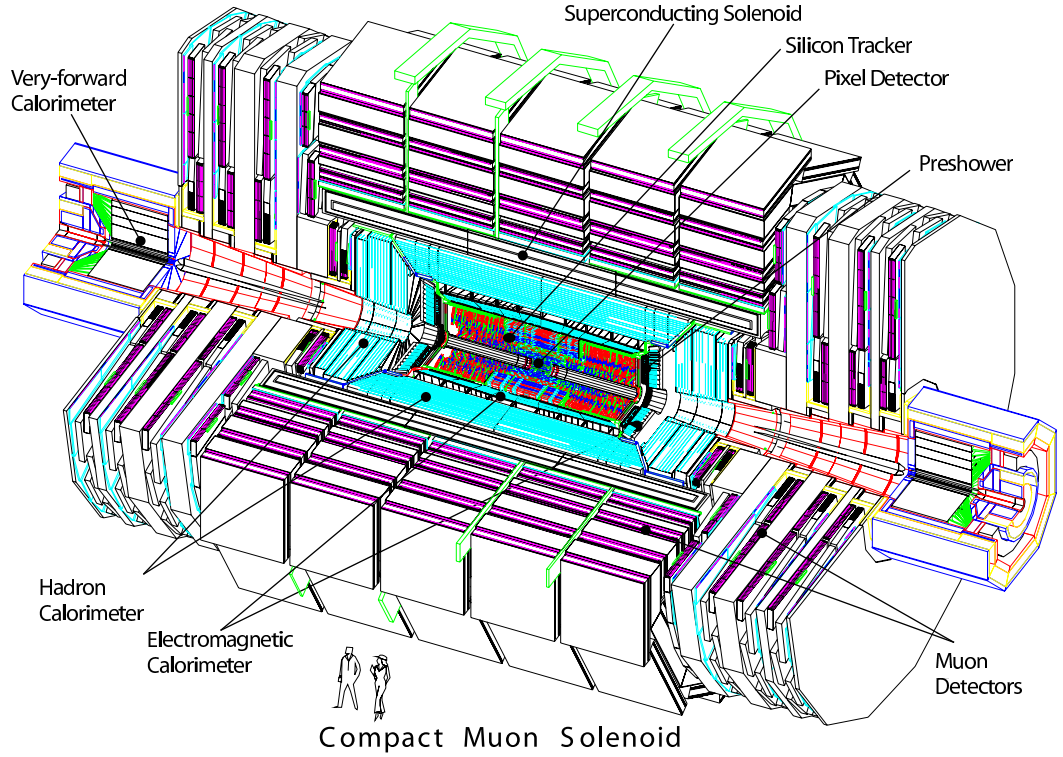

Figure 2.1: A schematic view of CMS detector, compared with human proportions. Sub-detectors are arranged in a typical concentric configuration, and the experiment is subdivided in a barrel part enclosed by two endcaps. Additional hadronic calorimeters are located in the very forward region.

2.2 The CMS detectors system

Figure 2.1 is a draw with a a complete layout of CMS. The CMS detector measures roughly 22 meters in length, 15 meters in diameter, and 12,500 tonnes in weight. It will reside on the 'Interaction Point 5' of the accelerator, sited in the French side (Cessy) in the fairest point from CERN. Most of the larger CMS part have been sub-assembled and fully commissioned in the experimental hall on the surface just above the cavern, than lowered piece-by-piece in their final destination, at $\sim$ 100 m depth.

The heart of the apparatus is a huge, high field (4 Tesla) solenoid, 13 meters in length and 5.9 meters in diameter. This magnetic field was chosen in order to achieve good momentum resolution within a compact spectrometer avoiding stringent demands on muon-chamber resolution and alignment. The return field is large enough to saturate 1.5 m of iron, allowing 4 muon "stations" to be integrated to ensure robustness and redundancy. The bore of the magnet coil is large enough to accommodate the electromagnetic and hadron calorimetry and the whole tracking system inside. Each muon station consists of several layers of aluminum drift tubes (DT) in the barrel region and cathode strip chambers (CSC) in the endcap region, complemented by resistive plate chambers (RPC).

The tracking volume is given by a cylinder of length 5.8 m and diameter 2.6 m. In order

to deal with high track multiplicities, CMS employs 10 layers of silicon microstrip detectors, which provide the required granularity and precision. In addition, 3 layers of silicon pixel detectors are placed close to the interaction region to improve the measurement of the impact parameter of charged-particle tracks, as well as the position of secondary vertexes.

The electromagnetic calorimeter (ECAL) uses lead tungstate ($PbWO_4$) crystals with a pseudorapidity coverage up to $|\eta|$ <3.0. The scintillation light is detected by silicon avalanche photo-diodes (APDs) in the barrel region and vacuum photo-triodes (VPTs) in the endcap region. A preshower system is installed in front of the endcap ECAL for the π^0 rejection.

The ECAL is surrounded by a brass/scintillator sampling hadron calorimeter with the same coverage. The scintillation light is converted by wavelength-shifting (WLS) fibers embedded in the scintillator tiles and channeled to photodetectors via clear fibers. This light is detected by novel photodetectors (hybrid photo-diodes, HPDs) that can provide gain and operate in high axial magnetic fields. This central calorimetry is supplemented by a "tail-catcher" in the barrel region, ensuring that hadronic showers are sampled with nearly 11 hadronic interaction lengths.

To improve detector coverage up to pseudorapidity of ±5.0, a hadron calorimeter consisting of iron and copper with embedded quartz fibers is added in the forward regions. The Cerenkov light emitted in the quartz fibers is detected by photomultipliers. The forward calorimeters ensure full geometric coverage for the measurement of the transverse energy in the event.

The thickness of the detector in radiation lengths is greater than 25 X_0 for the ECAL, and the thickness in interaction lengths varies from 7 to 11 λ_I for HCAL, depending on η.

All subsystems of CMS are bound by means of the data acquisition and trigger system, that aims to select events with high transverse momentum electrons, muons, photons or jets. It will be sketched in the next section.

2.2.1 The CMS trigger system

The LHC machine characteristics and the physics programme deeply influences the challenges of the first level trigger in CMS.

Differently from the majority of product from p-p collision, which have p_T lower than some GeVs, interesting physics with top quark is expected to be produced with large transverse momentum. High p_T event rate is dominated by jet production and this high QCD background does not allow hadronic-only final states to be selected in top decays. In addition, low-p_T muons from K and π decays as well as b and c quarks represent a large background to leptonic boson decays and a precise momentum measurement is therefore needed. Top events feature many jets, leptons and large missing transverse energy, thus isolation criteria, based on energy deposited around a cluster in the calorimeter or a track in the tracker, have to be used. Therefore, good selection capabilities for muons, electrons, jets and missing energy are mandatory for top analyses at LHC.

It is worth to stress that at the design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$, the LHC p-p collisions at the center of mass energy 14 TeV will deliver an event rate about 10^9 Hz, with an average of ∼25 "minimum bias" overlapping events. On the other hand, interesting events are only a small fraction of the total cross section: for instance, the rate of the non Standard Model processes addressed by this study will be found to not exceed some 10^{-6} Hz. Therefore, only 1 event out of $\approx 10^{15}$ has to be selected.

The Trigger/DAQ system accomplishes the task by managing the enormous number of read-out channels connected to the highly fine granularity of detectors. The bunch-crossing rate, designed to be equal to 40 MHz, is reduced on-line to approximately 100 Hz. The triggered events are stored on tape with a typical event size of 1-2 MB.

CMS experiment uses a multi-level trigger, which provide a quick rejection of background with a high and unbiased efficiency. An outline of the system follows.

The First Level CMS Trigger

First-level (L1) trigger must take a decision for each Bunch Crossing (BC) every 25 ns. Here trigger criteria are based on coarse information from clusters found in the calorimeters and from tracks in the muon trigger detectors, as well as some correlation of information between them. The event rate is reduced to about 100 kHz. Fast and simple selection algorithms are implemented in custom electronic.

The size of the LHC detectors and the underground caverns imposes a minimum transit time for signals from the front-end electronics, to reach the services cavern housing the trigger logic and distribute the decision back to the front-end. The total time allocated for the transit and for reaching a decision to keep or discard data from a particular beam crossing is 3.2 μs, where more than 2 μs are just for the transit. During this latency time, the detector data must be held in buffers and trigger processors have to work in pipelines, performing separate steps of the processing logic in parallel at each BC period.

A correct BC identification is crucial for event selection and for reducing the readout bandwidth. For this reason, the muon spectrometer are projected in order to allow a maximum time-of-flight comparable with the BC period, while the calorimeters requires that the pulse shape of the signals extends over many bunch crossings to be less sensitive to noise.

High trigger efficiency is ensured by the redundancy of selection criteria ("trigger menus"), which can concurrently select events of a wide range of physics studies and allow the measure of efficiency using data. Rates can be kept under control by adjusting thresholds of physical quantities or by a proper rescaling of objects with large cross sections.

In CMS, the Global Trigger (GT) is the top level of L1 Trigger System, that receives the 4 best candidates from the muon and calorimeters subsystem (with no isolation requirements). The trigger objects are called "trigger primitives" and could be photons, electrons, muons or jets. Specific quantities (as p_T, E_T, E_T^{miss}) are calculated and a 'rank' can be issued for each trigger primitives. Reduced-granularity and reduced-resolution data are used to form

primitives.

The highest rank objects arrive at the GT along with informations on their pseudorapidity and azimuthal angle (and charge for muons candidates). The GT then applies thresholds, event topologies conditions and other selection criteria according to physical or technical requirements. The GT logic can be programmed to calculate up to 128 different trigger menus in parallel for each BC. A final OR function combine all active algorithms and generates the L1 accept signal.

The High Level CMS Trigger

The High Level Trigger (HLT) of CMS provides a software-based event selection after the L1 hardware trigger.

The main feature of the CMS architecture is that HLT event selection is implemented as a system of algorithms running on large processor farm (Filter Farm). Avoiding any intermediate physical level in the selection chain, allows the HLT to be entirely software-implemented and to access to full resolution and granularity data, as well as calibration and alignment constants.

The data from each detector front-end belonging to a L1-accepted event are collected by a set of Read-out Units (where they can be compressed and zero-suppressed) and then delivered to the Builder Units (BUs), through a large switching network (Read-out Builder Network). The network bandwidth required is of the order of 1 TByte/s. The BU receiving the data fragments are responsible for the actual building of the event and, if requested by the specific HLT algorithms, serve it to a Filter Unit (FU) via another switching system, the Filter Farm Network. The FUs are the components of the Filter Farm where the HLT code is executed and the selections applied. Each processor runs the same HLT software code and reconstruction proceeds in the so-called 'virtual trigger' levels. In the virtual level-2 algorithms only calorimeters and muon informations are used, in the virtual level-3 tracker informations are included.

The use of a processor farm for all selections beyond Level-1 allows maximal benefit to be taken from the evolution of computing technology. Flexibility is maximized since there is complete freedom in the selection of the data to access, as well as in the sophistication of the algorithms.

Various strategies guide the development of the HLT code. Rather than exploit all possible data in an event all together, whenever possible only the objects belonging to limited regions of the detector that are actually needed are included ('regional' and 'local' reconstruction). The selected events are then forwarded to the computing services for storage or for further analysis.

The computing power needed by the Filter Farm corresponds to $\mathcal{O}(10^3)$ GHz processors. Given the event size of the order of 1 MB and the manageable output rate of $\mathcal{O}(10^2)$ Hz, 1 TByte of data will be written on mass storage every day.

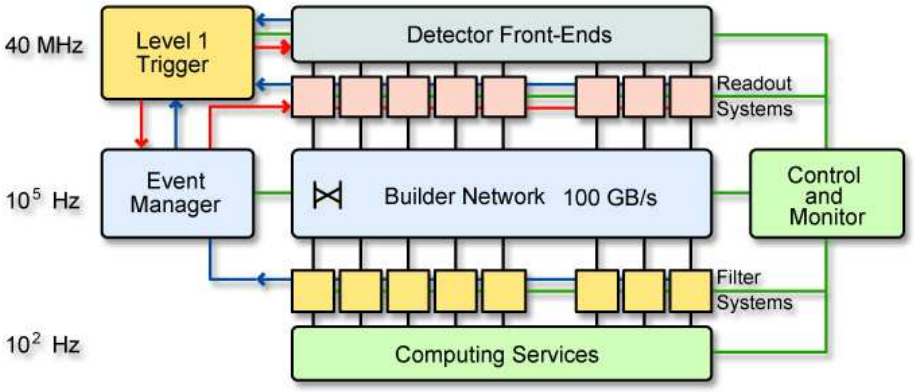

Figure 2.2: A sketch of CMS Trigger and Data Acquisition architecture.

Figure 2.2 is a picture of the trigger and data acquisition architecture, showing the detector electronics, the Level-1 trigger processors (calorimeter, muon, and global), the readout network, and an on-line event filter system (processor farm) that executes the software for the High-Level Triggers (HLT). Details can be found in Ref. [101] and Ref. [102].

2.3 Physics objects reconstruction in CMS

Once all sub-detectors will be fully commissioned and the trigger and data acquisition systems will prove to working properly, physics investigations will be ready to start. The task will be feasible only if physical objects will be reconstructed, employing as much information as possible from all the sub-detectors and combining them together. Some of the more common criteria used for objects reconstruction in CMS – and extensively exploited in the analysis described in the next chapter – are reminded below.

2.3.1 Photons and electrons reconstruction

The identification and reconstruction of electrons and photons with a p_T ranging from few GeV up to several hundreds GeV is mandatory for physics addressed in this work. The lower limit comes from the need to reject electrons from semileptonic decays of b quarks, the upper limit from W/Z boson decays produced by the top quark (and the photon produced in an anomalous decay), and the intermediate region includes leptonic decays of several background processes.

Both photons and electrons energy are measured in ECAL crystals, where they deposit all their energy. Electromagnetic energy is evaluated started from clusters built with arrays of 5×5 crystals (the "supercrystals"), that contain about 97% of the incident energy.

Before reaching the ECAL, a high-energy e or γ leaving the interaction point will cross all 10 layers of silicon tracker, as well as services. Material in front of the calorimeters results in photoconversion (for photons) and bremsstrahlung (for electrons). The material close to the

interaction point is the most dangerous, as the electron pairs or the electron and the radiated photon might be separated enough from the magnetic field in the tracking volume and originate separate clusters, resulting in a degradation of calorimeter performances. For example, electrons with $p_T = 35\,\mathrm{GeV}/c$ and $|\eta|$ =1.5 loose on average 43.6% in bremmstrahlung radiated energy. Part of this energy could be lost in the tracker material and never be recovered by the ECAL.

To minimize these effects on electromagnetic energy and position reconstruction, CMS has developed a topological cluster algorithm called "superclustering". The idea is to assume that bremsstrahlung radiation is spread by the strong magnetic field in a large area mostly along the ϕ direction, so the electron energy can be recovered by making 'clusters of clusters' along a ϕ road. In the ECAL barrel (EB), the superclustering ('Hybrid') algorithm dynamically searches for separated cluster in ϕ, exploiting in addition the lateral shower in η. In the ECAL endcap (EE), the 'Island' algorithm starts by searching for crystals with an energy above a threshold in ϕ and η. Along each scan line, crystals are added to the cluster until a rise in energy is encountered. In much the same way, non-overlapping clusters are grouped into superclusters, by collecting clusters in a very narrow η-window and much wider ϕ-window.

The e/γ energy is thus reconstructed by summing the energy deposited in a supercrystals or in a supercluster. Some specific corrections to the energy can be applied to recover energy lost inside the tracker or in cracks, and to account for variations of the shower containment as a function of the shower position.

A simple position measurement of the shower is obtained by calculating the mean position of the crystals in the cluster (centroid), usually weighted by a logarithmic function of crystal energy [103]. To increase hermeticity, the ECAL is designed so that the crystals do not point exactly to the nominal interaction vertex: therefore, the lateral position of the crystal axis depends on depth.

The resolution on the ratio between measured and real energy (reduced to a Gaussian distribution around 1) is 1.23% for electrons, 0.86% for unconverted and 1.15% for converted photons.

Photon reconstruction and selection

The off-line reconstruction and selection of prompt photons starts from a lower cut on the E_T of a cluster.

High energy electromagnetic showers could be originated also from high E_T neutral particles decaying into photons, such as π^0 and η, or directly from hadrons going toward the HCAL. Charged pions and kaons can be identified looking at the tracker, other fake photon signals due to jets can be reduced requiring isolation [104], by looking for additional energetic particles in a cone around the reconstructed ECAL cluster. The standard isolation variables considered can be based on the charged particle tracks reconstructed in the tracker with p_T above a given threshold, as:

- the sum of p_T for tracks in a cone ΔR around the ECAL cluster, or alternatively their number;

- the angle between the ECAL cluster and the nearest track direction.

It is also possible to consider only the tracks consistent with the primary vertex assigned to the event.

A different solution is to use energy deposits observed in the ECAL to discriminate single isolated photons from π^0 decays. Two variables are exploited:

- the sum of E_T from basic ECAL clusters, in a cone ΔR around the candidate supercluster or between two cones excluding the energy in the supercluster;

- the simple shower shape variable R9, defined as the fraction of the supercluster energy found inside the 3×3 array of crystals centered on the highest energy crystal.

A third handle to reject hadrons is a sum of hadronic energy deposits in the tower immediately behind the candidate supercluster. A proper combination of all these variables has been studied [102]. The Preshower detector in the endcap region, featuring a much finer granularity with respect to the ECAL, is exploited at this level to support the π^0 discrimination.

The photon direction is defined by a line between the event vertex and the centroid of the cluster. A precise determination of the photon direction relies on the capability to reconstruct photons originating from the primary interaction vertex that convert before impinging on ECAL. Figure 2.3 shows the integral distribution of the fraction of converted photons as a function of radius, measured in a fully simulated sample of $H \rightarrow \gamma\gamma$ events – the golden benchmark which dictated most of the ECAL properties. Conversion probability is displayed for 3 representative values of $|\eta|$: near the center of the barrel ($|\eta|$ =0.2), in the middle of a half barrel ($|\eta|$ =0.7), and toward the end of the barrel of the ECAL ($|\eta|$ =1.3).

A dedicated seed finding technique has been set up for tracking of the electrons from a converted photon [105]. Energy deposits in the ECAL are used as a starting point for a inward track seed finding. Then in the two outermost layers of the tracker, hits compatible with the trajectory hypothesis are sought (constraining for a maximum separation $\Delta\phi$ and a straight line from the primary vertex to the basic cluster). The inward tracking procedure results in a list of tracks of which only the opposite charge pair with the largest number of reconstructed hits is retained, whence, after cleaning and smoothing, parameters of the tracks at the vertex are obtained. When multiple two-tracks candidates are found for each supercluster, the one with the smallest invariant mass is chosen.

The reconstructed tracks originating from a γ can also be used at low luminosity to locate the primary vertex coordinate along the beam line. In the high luminosity scenario, this turns out to be much less easy and only the pointing information from calorimeters will be used.

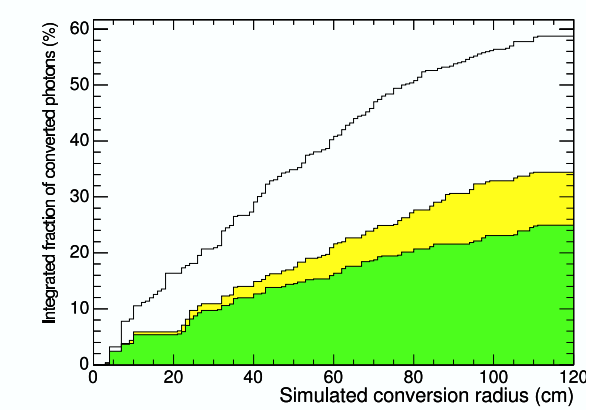

Figure 2.3: Fraction of converted photons from $H \to \gamma\gamma$ events, as a function of radius $r = \sqrt{x^2 + y^2}$. The 3 histograms correspond to slices of width $\Delta\eta = 0.1$ around $|\eta|$ =0.2 (dark grey), $|\eta|$ =0.7 (light grey) and $|\eta|$ =1.3 (empty histogram).

Electron reconstruction and selection

A primary electron is composed of a single track emerging from the interaction vertex and matched to an electromagnetic supercluster.

The main issue in electron reconstruction is the very large amount of bremsstrahlung emitted: about half of electrons of $p_T = 10\,\mathrm{GeV}/c$ (comparable to those from W coming from top decay) radiate 50% of their energy before reaching the surface of the ECAL. In about 10% of the cases, more than 95% of the initial electron energy is radiated.

The amplitude of the fluctuations in bremsstrahlung emission increases on average with increasing tracker material budget. This in general introduces largely non-Gaussian fluctuations, which affect the energy measurement in the ECAL and the momentum measurement in the tracker, as well as the electron identification observables.

For electron identification and reconstruction, the use of tracker information and cluster-track matching is the main tool. The superclusters drive the finding of the first 2 track hits in the pixel detector, then a match in $\Delta\phi$ and Δz to the innermost layer is applied. This technique allows an efficient separation of electrons from photons (that have a little chance to convert in pixel detector) and results in a 95% efficiency for isolated electron ($10 < p_T < 50\,\mathrm{GeV}/c$) with a rejection against QCD jet ∼17.

Thanks to the supercluster construction, in principle the energy-weighted average impact point of a non-radiating electron coincides with the one of an electron (of the same initial momentum) and its bremsstrahlung photons, as calculated using information from the supercluster in the ECAL. To build and fit the electron tracks across the whole Tracker, the two pixel hits found are exploited as seeds. This operation is performed by a non-linear filter approach such as the Gaussian Sum Filter (GSF, [106]), that generalizes the standard Kalman Filter [107] using a mixture of Gaussian distributions for all state vectors. The different components of the mixture measure different degrees of hardness of the bremsstrahlung.

Tracks from GSF are found to be quite sensitive to radiated bremsstrahlung energy, thus more suitable to describe the propagation of electrons than simpler approaches.

Distinct track-supercluster patterns (electron "classes") are recognized, exploiting the matching between the total energy collected by the supercluster with the momentum measured at the track origin (that is sensitive to the energy lost in the tracker material). Four classes [108] have been defined: *a*)"golden electrons", low-radiating objects with a reconstructed track well matching the supercluster, and a supercluster constituted by a single seed cluster; *b*) "big brem electrons", radiating more than a half of energy but without secondary photon conversion, thus with a good energy-momentum matching between the supercluster and the electron track at origin; *c*) "narrow electrons", with a bremsstrahlung merged in 1 cluster but a relaxed track-supercluster geometrical matching; *d*) "showering" electrons (more than 50% of the population) including *e.g.* cases of secondary conversion of a radiated photon, where a bad energy momentum matching is observed. The different classes influence the electron isolation and identification strategies.

The "golden" class contains about 20% of the selected electrons. They come mostly from W and Z decays, so they are the preferred objects for the off-line reconstruction detailed in the next chapter.

In addition, transverse shower profile and energy sharing between the calorimeters towers can be required to be consistent with a single electromagnetic shower, in order to suppress π^0 and multi-photon background. As a final step, the energy scale has to be corrected also for the energy lost in the tracker material, using special η-dependent factors based on the number of cluster.

The electron energy is finally determined by combining the corrected electron supercluster measurement, the electron track momentum and their associated errors. An energy resolution better than 4% can be obtained in the full energy range.

Electron and photon trigger

The electron trigger –as well as the muon one– will be a key issue in the analyses described in the next chapter, because it is the baseline choice to identify lepton from W and Z. The possibility of a (single) photon trigger will be also discussed.

The L1 trigger for e/γ objects is only based on E_T and isolation and foresees three different trigger streams: a single isolated trigger, a double isolated trigger and a double 'relaxed' trigger, with a slightly higher p_T threshold and the isolation requirement removed. Transverse energy thresholds delivering about 50% efficiency at low luminosity are respectively 23, 12 and 19 GeV . In the low luminosity phase, L1 trigger rates are 3400 Hz for the single isolated trigger, 1010 Hz for the double isolated and 420 Hz for the double non isolated [102].

The CMS HLT selection of electrons and photons proceeds in three steps.

At the virtual level-2 (L2) the candidate are reconstructed exploiting only the calorimetric information with the full granularity. The reconstruction is performed in the regions indicated

Table 2.3: HLT cuts for the electron streams.

Variable	Single electron	Double electron
$\|\eta\|$	< 2.5	< 2.5
p_T	$> 23\,\mathrm{GeV}/c$	$> 12\,\mathrm{GeV}/c$
ECAL Isolation	$< 3\,\mathrm{GeV}/c$	$< 9\,\mathrm{GeV}/c$
Track Isolation	< 0.06	< 0.4
E/p (Barrel)	< 1.50	–
E/p (Endcaps)	< 2.45	–

by the L1 candidates.

The second step (L2.5) demands hits in the pixel detector consistent with the L2 candidate. The expected hit position on the pixels layers is estimated by propagating inward the impact point of the candidate (weighted for the energy) to nominal vertex position. If at least 2 hits are found, the candidate is classified as an electron, otherwise as a photon.

In the final step (L3) the algorithm to select electron candidates has enough time to use the tracker hits in order to perform a full track finding and reconstruction. Electrons signals are identified adopting some additional criteria:

- matching between the energy in the ECAL (E) and the corresponding track momentum measured in the tracker (p) is imposed. In the single electron stream, an additional cut $E/p < 1.50(2.45)$ is applied in the barrel(endcaps), to reduce the hadron contamination of the stream;
- the most efficient isolation variables are found to be isolation in the electromagnetic calorimeter (a maximum energy of $3(9)\,\mathrm{GeV}$ in a cone $\Delta R < 0.15$ around the supercluster, for the single(double) electron stream), along with isolation in the tracker.

In the on-line reconstruction algorithms, a standard Kalman Filter technique is used, treating the electron as it were a muon, *i.e.* assuming a Bethe-Bloch distribution in the energy loss (both in the track building and in the fitting phase) and stopping the track when a significant amount of radiation have been emitted. Table 2.3 summarizes the high level trigger cuts for the two main electron streams.

The isolation requirements that are used at trigger level for the single photon are:

- the number of tracks with $p_T > 1.5\,\mathrm{GeV}/c$ inside a cone with $\Delta R < 0.3$ around the photon candidate less than 2;
- total E_T of all 'Island' basic clusters with $\Delta R < 0.3$ around the photon candidate less than $1.5\,\mathrm{GeV}$;

- the total transverse energy of HCAL towers within $\Delta R = 0.3$ less than 6(4) GeV in the barrel(endcaps).

The γ selection may also use lateral shower shape for π^0 rejection and reconstruction of converted photons. The rate of the photon candidates is further reduced applying higher threshold cuts than the electron stream, namely E_T >80 GeV in the HLT for the low luminosity scenario. Table 2.4 [102] contains the estimates for the single/double electron and the single/double photon, for the benchmark signals and from the QCD background.

Table 2.4: Output rates from HLT for electron and photon streams at the low luminosity phase.

	Signal		Background		Total
Single electron (p_T >26 GeV/c)	$W \to e\nu$	9.8 Hz	Jets	9.4 Hz	21 Hz
	$Z \to e^+e^-$	1.3 Hz			
Double electron (p_T^1, p_T^2 >12 GeV/c)	$Z \to e^+e^-$	1.1 Hz	Jets	0.8 Hz	1.8 Hz
Single photon (p_T >80 GeV/c)	γ + jet	2.1 Hz	Jets	1.4 Hz	3.5 Hz
Double photon (p_T^1 >30, p_T^2 >20 GeV/c)		≈0 Hz	Jets	1.9 Hz	2.3 Hz
			γ + jet	0.4 Hz	
Total:		14.3 Hz		13.9 Hz	28.2 Hz

2.3.2 Muons reconstruction and selection

As most of physics LHC is designed to explore, the signal addressed here is characterized by the signature of muons. Therefore, efficient muon identification with accurate momentum measurement is crucial to fully exploit the CMS physics potential, both in trigger and off-line reconstruction. To meet the LHC physics goals, a momentum resolution of about $10\%/p_T$[TeV] will be required.

Muon reconstruction steps

The muon reconstruction software is able to perform reconstruction in the muon system and the silicon tracker [102]. All algorithms are implemented using the concept of "regional reconstruction", that is the ability to reconstruct an object using the information from one or more limited region. Both the off-line reconstruction and the High-Level Trigger do not explore the entire tracker, but only the parts where a charged particle track compatible with hits in the muon chambers is likely to be found.

Muon reconstruction is performed in 3 stages:

Local reconstruction (local-pattern recognition). The first step of muon reconstruction is the reconstruction of the individual hits within both tracking detectors (DT and CSC) and RPCs, from where tracks segments are extracted;

Standalone reconstruction . The state vectors associated with the segments found in the innermost chambers are used to seed the muon trajectories, working from inside out using the Kalman Filter technique and solving the left-right ambiguity. In case no matching hits or segments are found (*e.g.* due to detector inefficiencies, geometrical cracks or hard showering) the search is continued in the next station. The track parameters and the corresponding errors are updated at each step, until the outermost surface of the muon system is reached. A backward Kalman Filter is then applied, working from outside in, and the track parameters are defined at the innermost muon station. Standalone muon reconstruction uses only information from the muon system and the precise information on the beam spot in the transverse plane ($\sigma \sim 20\,\mu$m);

Global reconstruction . Full muon tracks are then extended to include measurements of the silicon tracker. Starting from tracks parameters of the standalone reconstructed objects, the muon trajectory is extrapolated from the innermost station to the outer tracker surface, taking into account the effect of energy loss and multiple scattering. A Kalman Filter seeded by the hit pairs (both in pixel and microstrip detectors) is used for this purpose. The resulting trajectories are then matched to the original candidate in the muon system and their hits refitted together. The track-reconstruction algorithm consists of the following steps: trajectory building (that transforms each seed into a set of trajectories), trajectory cleaning (resolution of ambiguities between multiple trajectories) and trajectory smoothing (all reconstructed tracks are fitted once again).

The selected trajectories are then refitted excluding measurements (hits or segments) with high χ^2 values. The χ^2 probability of the fit is also compared with that of the tracker-only trajectory, in order to detect muon bremsstrahlung or any kind of significant energy loss before the first muon station. In the whole procedure, the energy deposit is required to be consistent with that of a minimum ionizing particle.

Muons off-line selection

The muon identification algorithm [109] provides an additional tool for selecting muon candidates, adopting an inside-out algorithm complementary to the standard muon reconstruction. The algorithm starts with all reconstructed tracks from the silicon tracking detector and attempts to quantify a muon compatibility for each of them (including tracks not matched with standalone tracks in the muon detectors). Muon identification takes advantage from pieces of information not used in reconstruction, such as associated energy deposits in the calorimeters and hits in the muon station not included in the global tracks.

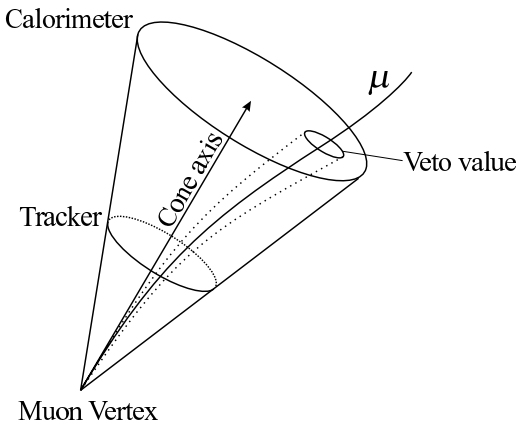

Figure 2.4: Schematic illustration of the isolation cone. The muon direction at the vertex defines the cone axis. The energy deposit in the cone is computed, and the muon contribution is removed by excluding the "veto value" from the cone.

The efficiency for reconstructing single, low-p_T muons is estimated to be 68.7%, that raises to 78.6% with the addition of muon identification algorithm. This algorithm is welcome also for the b-tagging method exploiting muons from b semileptonic decays ("soft μ-tagging", Sec. 2.3.4): the global reconstruction efficiency for these muons increases from a 71% to 84%.

As the last reconstruction criteria, isolation is the highest level variable on which high p_T studies commonly rely. In the momentum range relevant for triggering (p_T threshold in the range $10 \div 30\,\mathrm{GeV}/c$), the main sources of muons are b and c quark decays, along with low-p_T muons from muonic K and π decays. All of these muons are produced in usually soft jets and are thus accompanied by nearby particles. Isolation is a powerful tool to distinguish between the muons produced in jets and those coming from decays of heavy objects. The standard procedure is to compare of the total energy deposited in a cone around the muon (transverse energy in a calorimeter or the sum of transverse momenta of reconstructed tracks) with a predefined threshold. Figure 2.4 shows a schematic illustration of the isolation cone. The muon direction at the vertex defines the cone axis. The energy deposit in the cone is computed, and the contribution from the muon itself is removed by excluding the small area around the muon ("veto value") from the cone. Comparison of the deposit in a cone of given ΔR with a the muon energy determines the isolation.

First Level and High Level Muon Trigger

The First Level Muon Trigger of CMS uses all three kinds of muon detectors working independently. The excellent precision of DT and CSC ensures sharp momentum thresholds and their multilayer structure allows an efficient background rejection.

DT and CSC electronic processes firstly the informations in each chamber independently, delivering vectors (with position and angle) in any muon station. Vectors from different stations are collected by the Track Finder, which combines them to form a muon track, and assigns a transverse momentum value. Up to four muon candidates from each muon subsystem are selected and sent to the Global Muon Trigger. In the case of RPC there is not a local processing and, if the hits from different stations are aligned along a possible muon track, a p_T of track is assigned.

The on-line selection of muons within the HLT system [101, 102] is performed in 2 steps:

1. confirmation or discarding of the Level-1 decision, by using the Level-1 candidates as a seed and only information from the muon system;

2. application of p_T threshold (Level-2), reconstruction and use of the full tracker information (Level-3). Additional rate reductions can be achieved by requiring muons to be isolated in the calorimeter (after the Level-2) and in the tracker and pixel (Level-3).

For the low luminosity phase, an inclusive single-muon trigger rate of 25 Hz can be achieved with a p_T threshold of 19 GeV/c (including isolation). For di-muon triggers, a typical p_T threshold is 7 GeV/c, resulting in a rate of 4 Hz.

The step-by-step efficiency of the three-level algorithm is plotted in Fig. 2.5, where some 99% is obtained in almost the full pseudorapidiy range. For muons with $10< p_T <100$ GeV/c, momentum resolution is $\sigma(1/p_T) \sim 10\%$ in the barrel and $\sigma(1/p_T) \sim 16\%$ in the endcaps (p_T in GeV/c).

2.3.3 Jets and Missing energy reconstruction

Precise measurements of the jet energy and missing transverse energy is requested in most of physics CMS plans to address. In the present case, the need is to identify high energy jets (roughly $p_T >50$ GeV/c) and to tag the multi-jet background. It is worth to note that an accurate determination of the jet energy scale is a key ingredient to study the characteristic of the top quark.

Energy of hadrons and electrons is basically reconstructed by a weighted sum of deposits in ECAL (in a matrix around the central crystals) and deposits in HCAL (in a matrix around the central towers). Readout cells in HCAL are arranged in a tower pattern in (η, ϕ) space, projective to the nominal interaction point. Calorimeter towers are formed by addition of signals in η,ϕ bins corresponding to individual HCAL cells. The energy associated with a

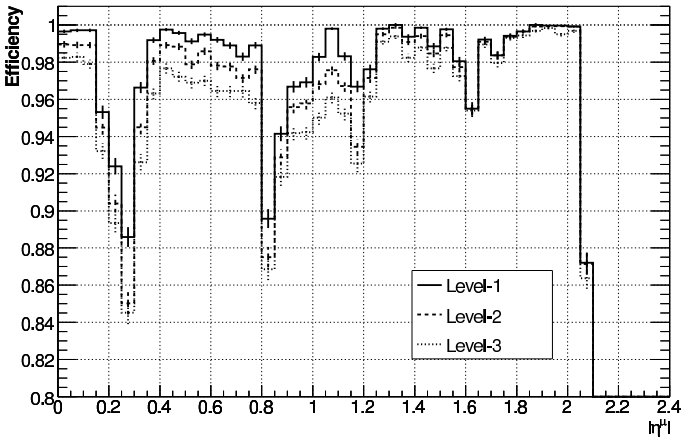

Figure 2.5: Cumulative efficiency for single muons to pass Level-1 (solid), Level-2 (dashed), and Level-3 (dot-dashed) trigger levels as a function of the generated muon pseudorapidity. No thresholds on p_T are applied here.

tower is calculated as the sum of all contributing readout cells which pass the on-line zero suppression threshold, along with any additional off-line software thresholds. Usually a cut E_T >0.5 GeV and E >0.8 GeV is applied to all towers used in jet reconstruction. Energy given by the tower energy, together with the direction defined by the interaction point and the center of the tower, is exploited to reconstruct the jets by different algorithms.

Jet reconstruction algorithms

A correct reconstruction algorithm has to associate the energy deposit in calorimeters cells to the energy of the final state partons produced in the hard interaction.

The jet reconstruction algorithm proceeds through two different steps. First the cell or cluster close to each other are associated, if they are near in angle (cone algorithm) or both in angle and transverse momentum (k_T algorithm), thus forming a jet. Subsequently, the kinematical quantities related to the jet itself (four-momentum) are calculated (*recombination scheme*) and used to compute η, ϕ and E_T.

The energy of the reconstructed jets is generally lower than those of particle belonging to the jet, due to two main reasons:

- detector effects: dead material and cracks, mis-calibration, non-Gaussian tails due to non-compensation, magnetic field effect, in-homogeneity of response due to different technologies in different $|\eta|$ regions, electronic noise;
- physics effects: initial and final state radiation, out-of-cone energy loss, longitudinal leakage for high p_T jets, underlying event and pile-up.

As a consequence, an absolute good precision on energy scale has to be reached and used

to calibrate jets.

Three principal jet reconstruction algorithms have been coded and studied for CMS: the iterative cone (the simplest and fastest, commonly used for jet reconstruction in software-based trigger systems), the mid-point cone and the inclusive k_T-jet algorithm (both widely used in off-line analysis).

Iterative cone algorithm (IC) . In the IC algorithm [110], calorimeter towers are sorted in decreasing E_T and a cone of size ΔR is cast around the input object having the largest transverse energy above a specified thresholds. Towers included in this cone are used to compute the E_T weighted centroid and jet direction, then a cone with a new radius ΔR is drawn around this direction. The procedure is iterated until the energy of this "proto"-jet changes by less than 1% and direction by $\Delta R < 0.01$ between iterations. When a stable proto-jet is found, all objects in the proto-jet are removed from the list of input objects and not used as input for another jet in the event. The whole procedure is repeated until the list contains no more objects with an E_T above the seed threshold;

Mid-point cone algorithm . In contrast to the IC algorithm, no object is removed from the input list [111]: this can result in overlapping proto-jets, when a single input object belongs to several proto-jets. For every pair of proto-jets that are closer than the cone diameter, a mid-point is calculated as the direction of the combined momentum. These mid-points are then used as additional seeds to find more proto-jets. When all proto-jets are found, the splitting and merging procedure is applied, starting with the highest E_T proto-jet. If the proto-jet does not share objects with other proto-jets, it is defined as a jet and removed from the proto-jet list. Otherwise, the shared transverse energy is compared with those of the highest E_T neighbor proto-jet. If the fraction is greater than a fraction (typically 50%) the proto-jets are merged, otherwise the shared objects are individually assigned to the proto-jet that is closest in the (η, ϕ) space. The procedure is repeated until no proto-jets are left;

k_T algorithm . The inclusive k_T jet algorithm [112] is a cluster-based jet algorithm, aimed to merge particles in order of increasing relative E_T, keeping their ΔR separation lower than some value. The cluster procedure starts with a list of input objects (stable particles or calorimeter cells). For each object i and each pair (i, j) the following distances are calculated:

$$\begin{aligned} d_i &= {E_T}_i^2 R^2, \\ d_{ij} &= min({E_T}_i^2, {E_T}_j^2) R_{ij}^2, \end{aligned} \tag{2.1}$$

where R^2 is a dimensionless parameter normally set to unity and $R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$. The algorithm searches for the smallest d_i or d_{ij}. If a value of type d_{ij} is the

smallest, the corresponding objects i and j are removed from the list of input objects and filled as one new object into the list of input objects. If a distance of type d_i is the smallest, then the corresponding object i is removed from the list of input objects and filled into the list of final jets. The procedure is repeated until all objects are included in jets. The algorithm successively merges objects which have a distance $R_{ij} < R$.

Jet algorithm resolution and linearity

Whatever jet clustering algorithm is chosen, the reconstructed jet energy E_T^{rec} must subsequently be rescaled to the parton energy E_T^{MC} obtained from the Monte Carlo simulation. An additional effect comes from the fact that the ECAL is calibrated on photons, whereas a substantial amount of jet energy deposited in the ECAL arises from charged pions.

A correction function can be obtained from a resolution plot, in which the width of a distribution of E_T^{rec}/E_T^{MC} is calculated. Then the width is fitted with the following functional form:

$$\sigma\left(\frac{E_T^{rec}}{E_T^{MC}}\right) / \left\langle\frac{E_T^{rec}}{E_T^{MC}}\right\rangle = \frac{a}{E_T^{MC}} \oplus \frac{b}{\sqrt{E_T^{MC}}} \oplus c \quad , \tag{2.2}$$

where the first term is due to energy fluctuations in the cone from electronics noise, pile-up and underlying event, the second term comes from the stochastic response of the calorimeter measurements and the last term comes from residual non-uniformities and non-linearities in the detector response. The symbol $\oplus$ stems for quadratic sum. Here a matching criterion based on the distance $\Delta R < 0.2$ is used to associate Monte Carlo particle-level and reconstructed jets.

Jet energy measurements (resolution, linearity, efficiencies, fakes etc.) are influenced by calorimeter response to both instrumental and physics-related effects. Due to these factors the average calorimeter response is not linear and depends upon the jet rapidity.

The resolutions of jet energy transverse E_T^{jet} quoted in literature [113], as resulting from a fit to Eq. 2.2 are:

$$\begin{aligned}
\sigma(E_T^{jet}, |\eta| < 1.4) &= 5.8\,\text{GeV} \oplus \frac{1.25[\text{GeV}^{1/2}]}{\sqrt{E_T^{jet}}} \oplus \frac{0.033[\text{GeV}]}{E_T^{jet}} \\
\sigma(E_T^{jet}, 1.4 < |\eta| < 3.0) &= 4.8\,\text{GeV} \oplus \frac{0.89[\text{GeV}^{1/2}]}{\sqrt{E_T^{jet}}} \oplus \frac{0.043[\text{GeV}]}{E_T^{jet}} \\
\sigma(E_T^{jet}, 3.0 < |\eta| < 5.0) &= 3.8\,\text{GeV} \oplus \frac{0.85[\text{GeV}]}{E_T^{jet}},
\end{aligned}$$

where the three η regions correspond to the barrel, endcap and forward HCAL.

Jet calibration

The scale factor between the real and reconstructed jet energy is generally known as *jet energy scale*; evaluating this scale is what is called 'calibration'. Therefore, uncertainties in jet energy are directly dependent from jet energy scale.

Jet calibration takes place in two steps:

1. the particle-level calibration corrects the reconstructed jet energy to equal the energy of particles in a jet from the hard scatter. The particle-level calibration has two components: offset and response. The offset component results from pile-up, underlying event and any residual electronic noise after calorimeter thresholds are applied, and has to be subtracted from the reconstructed jet. The response component results from non-linear response of the calorimeter to hadrons, and differences in response among the calorimeter regions in η;

2. the parton-level calibration corrects the energy of a particle-level corrected jet to the energy of the parton that originated the jet. The parton-level calibration accounts for the particles from the originating parton that are not included in the particle-level jet. This correction depends on the hadronization model used and on the type of originating parton.

The calibration technique described above is based on simulated samples and usually is the adopted tool in the analyses based on simulations – this one not excluded. Correction factors obtained depend on η and E_T and a set of these parameters should be provided for each jet finding algorithm, cone size, level of noise, pile-up. Going from simulations to real life, a careful calorimeter calibration has been undertaken with radioactive source and test beam, during all the assembling phase. Then from the very first LHC runs, real data (minimum bias or specific processes) can be exploited to cross-check the calibration. A few calibration methods are outlined below:

Data-driven calibration: The first stage of the calibration will check the radioactive source method at the tower level, measuring noise and adjusting thresholds, adjusting ϕ symmetry with minimum bias triggers, comparing tower-to-tower response using high p_T charged particle tracks. The second stage will check the calibration of jets by measuring the effect of pile-up on clustering algorithms and calibrating the jet energy scale;

di-jet balancing: QCD di-jet events can be employed to calibrate the jet energy scale vs. η and verify the resolution. The di-jet balance is defined as the ratio of the difference between p_T of a 'probe' jet and a jet in the barrel, and their average p_T. Di-jet balancing can be used to calibrations in different η regions ('inter-calibration'), based solely on the data. The jet resolution can be measured from the width of the di-jet balance, comparing the energy of reconstructed jets with that of partons at generator level;

photon+jet events: In the γ+jet events (in the limit of zero transverse momentum of the system) there is a direct photon with a p_T balanced by the jet. The high resolution of the electromagnetic calorimeter provides an accurate measurement of the photons and this reflects in a precise estimate of recoiling jet energy, nicely exceeding the HCAL one. It has been proposed [114] to statistically averaging over events with a fixed sum of photon and parton p_T (in order to include even the events with a momentum imbalance) and require tight cuts (transverse energy in the isolation cone $< 5\,\mathrm{GeV}$, angle between γ and jet $> 172^\circ$). This provides an absolute jet energy scale estimation (not only an inter-calibration), so allowing the systematic uncertainties due to initial state radiation to be estimated, as well as background from two jets with energetic π^0.

Missing transverse energy reconstruction

In top quark physics, the capability to measure the missing transverse energy with a good resolution is required, in order to reconstruct the W invariant mass when it decays in a lepton and a neutrino.

Assuming that the total transverse energy in an event is zero, the baseline strategy to evaluate E_T^{miss} is to catch all the visible transverse energy, measure its magnitude and direction and then perform a vectorial sum. The missing transverse energy will be this sum with the minus sign. The total energy deposit in the transverse plane can be thought to include two components [115]:

- the sum of jet energies p_T^{corr} over all the reconstructed jets, that have been clustered following one of the algorithm explained above. They are considered to be calibrated and included only if they have a p_T greater than a threshold (typically $20 \div 25\,\mathrm{GeV}/c$);
- all the contributions that can not be included in the previous class. This accounts for low-p_T jets and unclustered energy, and can be considered as the whole calorimeter energy E_T^{raw} subtracted by the sum of the uncalibrated jet energies p_T^{raw}. The quantity E_T^{raw} is obtained as the scalar sum of the transverse energy of all the raw calorimeter towers.

Therefore, a general expression for missing transverse energy is:

$$E_T^{miss} = -\left[\left(E_T^{raw} - \sum_{jets} p_T^{raw}(j)\right)_{\text{low } p_T} + \left(\sum_{jets} p_T^{corr}(j)\right)_{\text{high } p_T}\right] \quad , \tag{2.3}$$

where the two low/high p_T contributions are explicitly separated. A refinement of this algorithm includes the small energy deposit of the muons, that amounts to about $4\,\mathrm{GeV}$. As these low energies are much better measured by the tracker, this amount is replaced by the sum of muon p_T and added to equation 2.3.

In a machine as LHC, the presence of underlying event and pile-up collisions is an irreducible complication for the E_T^{miss} measurement. Since this measurement comes directly from the reconstruction of the energy and the position of jets, it will be further degraded by all the uncertainty sources in calorimeters as HCAL non-compensation, the difference between photon and pion response, the bending of tracks by the magnetic field, electronic noise and event synchronization, 'hot' or dead cells. The resolution of E_T^{miss} is considered as the distribution of the difference between the reconstructed E_T^{miss} and the transverse energy of non-visible particles from simulated event samples. It is proportional to $\sqrt{E_T^{raw}}$, because jet energy resolution follows this rule. The resolution of the E_T^{miss} in QCD di-jet events with pile-up is given by $\sigma(E_T^{miss}) \approx 1.0\sqrt{\Sigma E_T}$ if energy clustering corrections are not made, while the average E_T^{miss} is given by $E_T^{miss} \approx 1.25\sqrt{\Sigma E_T}$.

First Level Calorimeter Trigger

Trigger on jets is a common practice in many analyses, mainly those oriented to SUSY. The transverse energy sum is computed for each calorimeter trigger tower, using different energy scales in different zones of the calorimeters. Programmable thresholds can be applied to individuals calorimeter channels before the trigger primitives calculation, in order to improve the immunity to noise.

In each calorimeter region, defined by 4×4 trigger towers, the highest E_T isolated and non-isolated jet candidates are separately found. The four highest E_T candidates of both categories in a $\Delta\eta \times \Delta\phi = 3.0 \times 0.7$ region are transferred to the Global Calorimeter Trigger, where the top four candidates are retained for processing by the CMS GT.

The jet trigger uses the transverse energy sum computed in the calorimeter regions extending up to $|\eta|$ =5.0 (except in HF where it is computed on the tower itself) and the corresponding trigger is defined by a threshold value and a pre-scaling factor.

2.3.4 Beauty selection in CMS

Many interesting physics channels produce b-jets in the final state. The Standard Model top quark is one among the main sources of these objects. Therefore, detect the b quark and distinguish it from a copious backgrounds containing only light flavoured jets is crucial for the reconstruction of the SM top decay.

Inclusive tagging of b-jets relies upon properties that are distinctive of b-hadrons: large proper lifetime ($\tau \approx 1.5\,\text{ps}$, $c\tau \approx 450\,\mu m$), large mass, decays to final states with high charged particle track multiplicities (on average 5), relatively large semileptonic branching ratios (in about 20% of the cases, b-hadrons decay into muons or electrons).

As a consequence, several techniques had been envisaged for b-tagging, as study of lepton energy in b-quark semileptonic decays, impact parameter significance, track multiplicity at secondary vertex, flight distance etc. In CMS, reconstruction algorithms use much of this

information and produce a fit, with a χ^2 value that can be used for a partial discrimination between jets with and without beauty.

Most of the b-hadron properties used for b-tagging are exploited using charged particle tracks, so track reconstruction has to be precise enough. The following track selection cuts are common to all b-tagging algorithms described here: *a*) at least 8 reconstructed hits in total, among which at least 2 reconstructed hits in the pixel detectors; *b*) transverse momentum $p_T > 1\,\mathrm{GeV}$; *c*) χ^2/ndf of the track fit < 10; *d*) transverse impact parameter with respect to the reconstructed primary vertex < 2 mm. This serves to reject charged particle tracks originating from sources as Λ^0 and K^0 (the so-called 'V^0 decays'), photon conversions and nuclear interactions (in the beam-pipe or the first layers of the pixel detector).

The efficiency ϵ_q to tag a certain flavour q of jet as a b-jet (b-tagging efficiency for b-jets, misidentification or "mistagging" efficiency for non-b jets) is defined as the ratio between the number of jets of flavour q tagged as b and the number of jets of flavour q, determined by analyzing the parton content in a cone around the jet direction. It has to be taken into account that a gluon jet has a non-negligible probability to split into a $b\bar{b}$ pair (with a rate typically at the level of a few percent, depending on the energy of the gluon).

In the following, the algorithms currently exploited in CMS for b-tagging are briefly outlined. They can be applied both off-line and at the level of the High-Level Trigger [101].

Track impact parameter based tags

The main feature that distinguish the jets coming from a b quark from those coming from a light quark is that they contain tracks with a large impact parameters, originating from a secondary vertex. In fact, the lifetime of hadrons with beauty $\sim$1.5 ps corresponds to a decay length of around half millimeter in the rest frame.

As a result, the track impact parameters of charged particles is a powerful indication of the distance traveled by a particle from the primary interaction vertex. The impact parameters distribution has a 'bulk' centered on zero, (that originates from finite detector resolution, badly reconstructed primary vertex positions, fake tracks and multiple scattering of particles in material at small radii) and a 'tail' to the right, associated to a genuine track coming from a decay at a point with some positive displacement from the primary vertex.

The ratio (with sign) of the track impact parameter to its experimental uncertainty is commonly known as the 'impact parameter significance' and it is a powerful discriminating variable against non-b-jets. In the track counting b-tagging method, the impact parameter significance is computed for each selected track in a jet. The jet is tagged if the distance of closest approach to the jet axis is less than 0.07 cm and if the number of tracks with an impact parameter significance exceeding a given cut is greater than a given value, which can be optimized for various purposes. The "discriminator" for this algorithm is the impact parameter significance.

Probability b-tagging

The probability method [116] entails computing the compatibility of a set of tracks with having come from the primary vertex.

For each track, the probability of coming from the primary vertex is computed, then all probabilities are combined together to provide the 'jet probability'. The track probability is computed using, as calibration, the negative part of the distribution of the impact parameter significance: a "resolution function" is determined from it and a probability $P(x)$ of a track having impact parameter significance x is extracted. The track reconstruction quality is related to the momentum and the number of hits in the different types of detectors involved. Here the discriminator is related to the jet probability of a jet containing N tracks, defined as the confidence level for a group of N tracks to produce the observed, or any less likely value of track probability. The distribution of this parameter is close to zero for light quark jets and has a wider distribution for b-jets, providing a very good separation.

Combined secondary vertex tag

The combined b-tagging algorithm [117] is based on the reconstruction of the secondary decay vertex of a weakly decaying b-hadron: the presence of a secondary vertex likely indicates the decay of a hadron with beauty.

Different topological and kinematic vertex variables are combined together with track impact parameter significances into a discriminating variable to distinguish b-quark jets from non-b-jets.

The following cuts are applied to the resulting vertexes to select secondary vertex candidates:

- the distance from the primary vertex to the secondary vertex in the transverse plane has to exceed $100\,\mu$m and must not exceed 2.5 cm;
- the distance from the primary vertex to the secondary vertex in the transverse plane divided by its error has to be greater than 3;
- the invariant mass of charged particles associated to the vertex must not exceed 6.5 GeV ;
- the vertex must not be compatible with a V^0 decay.

Depending from the output of these requirements, additional variables as the energy fraction and the rapidity of charged particles associated to the secondary vertex are considered. All accepted track impact parameter significances enter into the final discriminator. In order to improve the suppression of charm quark jets, a lower threshold to the mass of selected tracks is added to the discriminator. Thus a Likelihood function $\mathcal{L}^{b,c,q}$ that measures the probability for a jet to be a b-, c- or $udsg$-jet can be defined. A 'combined' discriminator $Disc$ is given by:

$$Disc = f(c) \times \frac{\mathcal{L}^b}{\mathcal{L}^b + \mathcal{L}^c} + f(udsg) \times \frac{\mathcal{L}^b}{\mathcal{L}^b + \mathcal{L}^{udsg}} \quad ,$$

where $f(c)$, $f(udsg)$ are the expected content of quarks and gluon content. This parameter provides a good discrimination between b- and $udsg$- jets (for $Disc$ above ~ 0.5) and between b- and c- jets (for $Disc$ above ~ 0.9).

Soft lepton tags

Finally, the soft lepton b-tagging algorithm [118] is based on the relatively high b quark branching ratio to electrons and muons, which is about 19% for each lepton family.

The key element required to take advantage of this property is the identification of leptons among the tracks to each jet. To increase the purity of the selection, additional cuts are applied to parameters associated with these leptons. Muon identification is performed looking for reconstructed muons in a region around the jet axis, which is the best approximation to the original quark direction. Electron identification uses the track with $p_T > 2\,\mathrm{GeV}/c$ and $|\eta| < 1.2$ matched with ECAL clusters, then several variables (such as the covariance of cluster energy distribution, repartition of cluster energy and cluster energy and track momentum ratio) are combined into a single discriminating variable using a neural network technique. The topology of lepton tracks and jets (distance in the $\eta - \phi$ plane, energy ratio, track momentum and impact parameter) are exploited to feed a second neural-network, whose parameters are optimized for the best efficiency and purity.

2.3.5 Detector simulation and reconstruction tools

The reconstruction algorithms described above, as well as the simulation tools adopted to study the performance of these algorithms, have been implemented in the object-oriented framework COBRA [119]. The collection of detector simulation and reconstruction code developed in this framework is known as ORCA [94], and with COBRA has served the simulation and design needs of the Collaboration.

Full simulation for CMS detector

The simulation of detector geometry, materials and magnetic field is a package called OSCAR [120], that is implemented for all CMS sub-detectors in both the central region (Tracker, Calorimeters and Muon Systems) and in the forward regions (CASTOR calorimeter, TOTEM telescopes, Roman Pot detectors and the Zero Degree Calorimeter). OSCAR reads the individual generated events and simulates the effects of energy loss, multiple scattering and showering in the detector materials, relying on the object-oriented toolkit GEANT 4 [121]. GEANT 4 provides a rich set of physics processes describing electromagnetic and hadronic interactions in detail. It also provides tools for modeling the full CMS detector geometry

and the interfaces required for retrieving information from particle tracking through these detectors and the magnetic field.

The ORCA package takes care as well of the digitization (simulation of the electronic response), the emulation of the Level-1 and High-Level Triggers and the off-line reconstruction of physics objects. Each operation is aimed to resemble as much as possible all what will be set-up during the LHC operation.

Once all the hits in the detector are simulated and digitized, ORCA delivers a series of "high-level objects" as output. These quantities, such as reconstructed hits for charged particles in the tracker layers or energy deposits in calorimeter cells, can then be used as inputs of the higher-level algorithms as track fitting, calorimeter clustering, b-tagging, electron identification, trigger algorithms, jet reconstruction and calibration. After this step, objects on which analysis can be performed are at hand.

Analysis with the ROOT framework

The analyses contained in the next chapters have been possible thanks to a tool available to the CMS Collaboration, called `ExRootAnalysis`.

It is a package designed to collect all the objects delivered by the simulation and reconstruction framework, and organize them in such a way to be easily handled by ROOT [122], the standard data-handling framework for high energy physics.

`ExRootAnalysis` can be used to interface directly the file with reconstructed events. The objects the user wants to reconstruct can be set simply by switching the proper modules, possibly choosing the options for the implemented algorithm. The program has an access to the GEANT related information (simulated hits, tracks and vertexes), L1 and HLT trigger bits, basic reconstructed objects (calorimeter clusters, track segments etc.) etc., thus for each event it can take the off-line candidates (electrons, muons, jets etc.) and evaluate the high-level variables, *e.g.* energy, position, number of track for the jets, E/p for the electrons, discriminator for the b-jets and so on. For a given number of events, these variables are organized as ROOT-histograms and coded as ROOT-tree branches, one for each class of high-level reconstructed objects (or more, if different options are activated). These classes will be simply called and handled by the user analysis program, that at this level can be written as a common ROOT macro.

Such an approach will be followed in all the analysis described hereafter.

Fast simulation for CMS detector

A framework for fast simulation of particle interactions in the CMS detector has also been developed by the Collaboration. It is called FAMOS (FAst MOnteCarlo Simulation, [123]) and has been intended to be used for preliminary physics analysis, development of reconstruction algorithms, generation of very large samples, switching of generation parameters to study

theoretical systematic effects and, in general, all applications where computing time could be an issue. It is an object-oriented system for which C++ has been chosen as programming language.

The input of FAMOS is a list of particles (originating from an event generator or a particle gun) characterized by their momentum and origin vertex, with mother and daughter relationships following the various decay chains in the event. Upon user request, each of the (quasi)-stable particles in this list is then propagated in the CMS magnetic field to the different layers of the sub-detectors, which it may interact with. While propagating, these particles are allowed to decay according to their known branching fractions and decay kinematics. The particles resulting from the interactions with the detector layers or from the decays in flight are added to the original list, and propagated/decayed in the same way. The interactions simulated in FAMOS are: *a*) electron bremsstrahlung; *b*) photon conversion; *c*) charged particle energy loss by ionization; *d*) charged particle multiple scattering; and *e*) electron, photon, and hadron showering. The first four are applied to particles traversing the thin layers of the tracker while the latter is parameterized in the electromagnetic and hadron calorimeters. Most of simulations are based on a parameterization of efficiencies and resolutions of calorimeters energy, electron and muon momentum, along with the effect of the magnetic field on jet reconstruction.

The baseline choices implemented in FAMOS can be briefly sketched in the following way:

Tracks reconstruction for charged particles is implemented with a combinatorial track finder, that matches the objects reconstructed in the tracker with clusters in calorimeters. Identification of isolated muon and electrons is always possible;

Jets can be built with Iterative Cone or k_T algorithm, and different methods are implemented for missing transverse energy evaluation. Rescaling to obtain parton energy from calorimetric towers is done with constant coefficients (depending from E_T and η) deduced from dedicated Monte Carlo simulations;

b-tagging clones the one used in the full simulation (that uses a combined b-tag) and producing the "discriminator value" that represent a measure of the probability for a jet to be a b-jet;

photon and electron reconstruction implement a calibration algorithm (that takes account of energy losses, rear and lateral leakage, bremsstrahlung, photo-conversion etc.) in order to find the energy and position of electromagnetic particles with the best accuracy.

The computer time needed to simulate an event in FAMOS is about 3 orders of magnitude smaller than that needed in the full chain, and the level of agreement is at the percent or below.

With the only exception of the HLT, the high-level objects are provided to the user in the same way as in the full reconstruction and analysis package. This parallelism between the fast

simulation and the complete reconstruction has several advantages: among them, it allows the managing of the sample generated with the full simulations in the same way as those generated with the fast framework. As a consequence, switching between analysis programs that handle the objects delivered by the two packages do not require major modifications. A clone of the `ExRootAnalysis` service is available for interfacing to FAMOS objects.

Pile-Up Treatment

The total inelastic cross section at the LHC is assumed to be 80 mb. The LHC will operate at a bunch crossing rate of 40 MHz but 80% of the bunches will be filled, resulting in an effective bunch crossing rate of 32 MHz.

During all its operation time, the LHC accelerator will produce a number of inelastic ('hard-core') pp collisions that will "pile-up" on top of the signal collision firing the trigger. The distribution of the number of these events is a Poissonian with a mean value μ depending from the instantaneous luminosity: in the low luminosity ($\mathcal{L} = 2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$) and high luminosity ($\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$) phases of operation. If also diffractive collisions are considered, the total pile-up increases to 5.0 and 25 collisions for low and high luminosity operation.

Both the simulation and reconstruction chain OSCAR/ORCA and FAMOS allow the overlay of Poissonian pile-up events, on top of real signal events exactly as for real data. Since the addition of pile-up collisions occurs much faster than the detector simulation, pile-up collisions are simulated separately from the signal collisions. Both outputs are merged in a second step, using a luminosity dependent pile-up contribution. Generated signal collisions are then re-used for producing samples corresponding to different luminosities.

Chapter 3

The analysis of two top FCNC channels

The present chapter is the main of the work, because it outlines the strategies that have been adopted to detect the FCNC top decay in presence of a Standard Model background. The addressed new physics channels are two: the neutral decay of top into Z, $t \to Zq$ and the one into a photon, $t \to \gamma q$. In both cases, top quark is produced in association with an anti-top following the standard decay, since the identification of a $t \to Wb$ process is a very powerful mean to flag the existence of an anti-top recoiling against the first. All the analyses are based on a set of properly optimized cuts.

The first section of the chapter describes the topology of the signal, that determines the peculiarities that made it discernible from the background – the so-called "signature". The signature dictates the kinds of background processes that are likely to hinder the signal identification. A detailed list of background for the two channels follows, along with the cross-section times the relevant branching ratio, that is proportional to the background rate so it expresses how important a background source is. The generator programs and parameters that have been used are briefly outlined.

Section 3.2 is preliminar to the analysis and describes the toolkit that made possible the selection procedure. The choices that are common to both the analyses are motivated in this part. Namely, the parameter adopted for the b-tagging and for selection of non-b-jet are optimized in such a way that maximizes the reconstruction efficiency and maintains the mis-identifications to a manageable level.

The analysis if the $t \to qZ$ decay is detailed in Sec. 3.3. An extensive discussion of how electrons and muons (that are chosen by the trigger stream) can be best reconstructed, and which reconstruction performance can be obtained is issued. Once the most important background sources have been identified by an *a posteriori* observation, efforts are put to devise a selection capable to reduce the impact of these sources. A set of optimized cut is thus established. With these optimized cuts, high level objects can be reconstructed, as the Z

and the W boson, the top quark decaying in the standard way and that following the FCNC coupling. The final result is that, retaining an efficiency of some percent, all the background sources can be maintained at a very low level.

The study of the $t \to \gamma q$ decay, contained in Sec. 3.4, begins with a discussion of the best trigger to choose. Performance of a standard selection on muon, electrons and photons resembling the previous case are presented, then they are compared with the choice of a single-photon trigger stream. The conclusion is that, with a specific set of pre-selection cuts, the photon trigger has slightly better performance and should be preferred. Anomalous decays can be successfully identified even in this case and background suppression is equally strong.

In the last section, the detailed results of the two analyses are considered, to check if other processes that have not been generated could have some chances to survive the selections. These processes may include both the ones expected from the Standard Model, which cross section and dynamic features have been calculated, and new physics process, that LHC is supposed to reveal. The conclusion is drawn that, under the reasonable assumption that rates for the new physics events are much smaller than those for expected ones, no processes that have not been simulated should impact the final results in a sizable way.

3.1 Signal and background events generation

3.1.1 The FCNC signal

The signal that has been chosen to be studied is made up of a $t\bar{t}$ pair, produced in the p-p interaction via an exclusive process. One of the top quark is assumed to decay in the SM final state, $t \to W^+ b$ ($\bar{t} \to W^- \bar{b}$) and the other to decay through the FCNC process, $\bar{t} \to V\bar{q}$ ($t \to Vq$), where $q = u, c$. This latter final state marks the difference with the common $t\bar{t}$ production followed by the standard decay. The aim of the study is to determine which are the sensitivity and discovery reach of CMS toward this channel. It will be accomplished by counting the minimum number of events that are visible on top of Standard Model background.

The whole analysis is based on a cascade of selection algorithms, aiming to obtain a strong background reduction while retaining a good signal efficiency. Therefore, choosing a specific decay mode that is expected to have a distinctive experimental signature is an important prerequisite. In this view, decay channels leading to a rich hadronic final state strongly hinder the signal discovery, and pieces of information they could add to the final result are supposed to have low significance. Therefore, two limitations are imposed to the simulated signal channel:

1. the $t \to Vq$ is restricted to $V = \gamma, Z^0$. As explained in the first chapter, limits on the tgq coupling are best suited to modes different from the $t \to gq$ decay;

2. the leptonic channels in electrons and muons are the only considered for the decays of

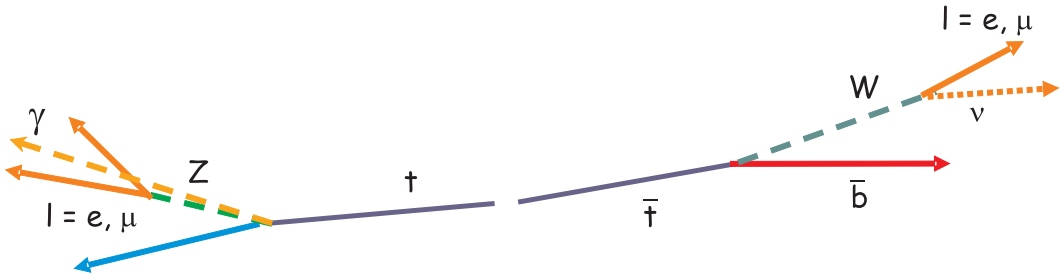

Figure 3.1: An picture of the event topology addressed in the work, in which one top quark decays to $Z/\gamma q$ and the other to Wb.

both Z^0 and $W^{\pm}$.

These restrictions, though seeming hard, are essential to cope with the large multi-jet background and set a signal with a very distinctive experimental signature. Figure 3.1 is a picture of this specific partonic final state.

The signal signature

The first step in the approach to the analysis is to identify the relevant features in the signal topology. This is the starting point to search for background processes that are likely to occur at LHC energy and luminosity, and may hinder the signal reconstruction. As a consequence, the signal signature draws the guidelines to devise the optimal selection strategy. The features that mostly make different this signal form the background events are sketched below:

- isolated, high transverse momentum charged leptons ($p_T > 10 \div 20\,\mathrm{GeV}/c$) coming from vectorial boson decays. From these leptons, reconstruction of W (transverse) mass and Z^0 (for the $t \to qZ$ case) is possible;
- in the $t \to q\gamma$ case, a high-energy isolated photon;
- large missing energy ($E_T^{miss} > 30 \div 40\,\mathrm{GeV}$) from undetected neutrinos in W decays;
- two hard jets, typically with transverse momentum $p_T > 50 \div 60\,\mathrm{GeV}/c$, coming from the fragmentation of b and q quarks. In the top frame of reference, quark coming from top have a $p_T = (m_t/2)\left(1 - M_W^2/m_t^2\right) \simeq 65\,\mathrm{GeV}/c$;
- one of the jets identified as containing a b quark, with some b-tagging algorithm.

3.1.2 The relevant background

Once the signal topology has been defined, the second step is to take into account all the processes that share some features with the signal, thus having a chance to enter the selections. Such processes are basically the reactions the Standard Model foresees in the p-p interaction

at $\sqrt{s}$=14 TeV, leading to a final state that could not be fully rejected. The impact of background processes on the analysis is proportional to their cross sections, which are calculable from the SM in most cases up to the Next-to-Leading-Order (NLO) – if not better. In the whole work, only the kinematic region with high transverse momentum ($p_T > 5 \div 10\,\mathrm{GeV}/c$) and central ($|\eta| < 2.5$) objects is taken into account, as most of the interesting physics is supposed to lie there.

In the following, a description about the possible background sources is outlined, separately for the two analyzed signals. Current values about the most up-to-date cross sections are quoted and referred. With the choice of the signal decay channels outlined above, no 'irreducible' backgrounds are present, *i.e.* processes whose final state is indistinguishable from the signal are naturally ruled out. The 'reducible' background shares only some features with the signal, thus it can be effectively swept-away by the selection cascade. Due to the presence of 'fake' objects and impurities in the selected samples, part of background survives till the last selection, and has to be subtracted from the final amount of event.

Adopting the common jargon, the scale factor between the NLO and LO cross section for a specific process is referred as a 'k-factor', where $k_{(NLO)} = \sigma_{NLO}/\sigma_{LO}$. Usually the quoted ratio comes from cross sections integrated in the whole kinematic region, although the difference in the magnitude may depend from several variables (as p_T of the final state particles, or the angular distributions or the unphysical factorization scale). The next chapter, in which an estimation of background is proposed from suitable data samples, will shown that a detailed knowledge of these dependencies is not relevant for the scope. Consequently, here k-factors are introduced mostly as a simple overall scale.

Except when explicitly stated, in all the background the following kinematical cuts in generation are adopted for leptons (l), jets (j) and invariant masses ($M(ll)$, $M(jj)$, $M(b\bar{b})$):

- jets are defined by the k_T algorithm (see Sec. 2.3.3) in $\Delta R < 0.4$ cone, having $p_T(j) > 20\,\mathrm{GeV}/c$, $|\eta| < 4.5$ and separation from leptons $\Delta R_{lj} > 0.4$;
- leptons have $p_T(l) > 15\,\mathrm{GeV}/c$, $|\eta| < 2.5$, separation $\Delta R_{ll} > 0.2$ and $M(l^+l^-) > 15\,\mathrm{GeV}/c^2$, to avoid the abundant Drell-Yan production by virtual photons in the Z/γ^* processes;
- for b-jets, $|\eta| < 2.5$ and $M(b\bar{b}) > 10\,\mathrm{GeV}/c^2$.

Bottom quarks are taken to be massless, as the massive results are not yet known at NLO, and processes in which there are b quarks already present in the initial state are ignored.

Finally, only the Standard Model processes are considered here. Some new physics processes that could mix with the FCNC signal will be addressed at the very end of the chapter.

Backgrounds for the $t \to Zq$ channel

The most peculiar features of the $t \to Zq$ channel are the presence of large missing energy (with a median about $50 \div 60\,\mathrm{GeV}$ and a tail up to $250\,\mathrm{GeV}$) and three leptons with high

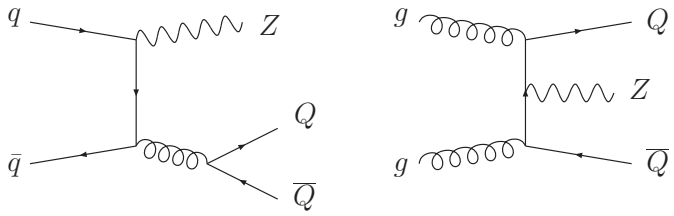

Figure 3.2: Diagrams contributing to the associated production of a Z boson and two high-p_T b quark. Due to the high gluon density in p-p collisions at 14 TeV, this process is quite abundant at LHC. Cross-section of this process strongly diverges with increasing $\hat{p}_T$.

transverse momentum (two same-sign and one different). In addition, the events must contain two hard jets, coming from the fragmentation of the b quark from the SM top decay and the light quark from the FCNC top decay.

In principle, two classes of processes may emulate all these signal features, thus they have to be taken into account:

- the ones producing three hard isolated leptons, in addition with jets coming from QCD processes (*e.g.* from flavour excitations, initial/final state QCD radiation, gluon splitting etc.);
- the ones producing two hard jets or more (with some of them originating from b quark), in addition with one or more hard lepton (*e.g.* from leptonic decays of heavy flavours in jets).

The 'reducible' background sources which have been considered are listed and discussed below:

- *$t\bar{t} \to (bl^+\nu)(\bar{b}l^-\bar{\nu})$ production*: The final state most similar to the signal is issued by the $t\bar{t}$ production, followed from the SM decay $t \to bW$ and a leptonic boson decay $W \to l\nu$ ($l = e, \mu, \tau$). While the SM side is indistinguishable from the signal one, the Z^0 from the FCNC side can be faked by the lepton from W and a third lepton (oppositely charged), from the semileptonic b quark decay, or from some hadron misidentified as a lepton. If in addition the identification of one b-jet fails (interpreting it as a light-jet), the final configuration of the signal is fully imitated. As it will be shown these mis-identifications are not unlikely to occur, hence this is largely the main background source. Taking the $\sigma(t\bar{t})$ quoted in the first chapter and $BR(W \to l\nu)$=0.327, the cross section of the process is 825 pb $\times BR(W \to l\nu)^2 = 88.44$ pb ($l = e, \mu, \tau$);
- *Vectorial boson production in association with a $b\bar{b}$ pair*: The associate production of a $b\bar{b}$ couple, together with a W/Z boson is a common process at LHC and may occur via both gg and $q\bar{q}'$ interaction. Figure 3.2 offers the lowest order Feynman diagrams for this production. The hardness of the interaction is usually measured by $\hat{p}_T$, the

Table 3.1: The NLO cross-section values (in pb) for the $pp \to Z(\to l^+l^-) + b\bar{b}$ production, in two of the relevant $M(l^+l^-)$ bins. The k-factors have been estimated with the MCFM [124].

	$60 < M(l^+l^-) < 100\,\mathrm{GeV}/c^2$	for $M(l^+l^-) > 100\,\mathrm{GeV}/c^2$
$\sigma(Z \to (ee) + b\bar{b})$	26.206 ± 0.028	1.0 ± 0.1
$\sigma(Z \to (\mu\mu) + b\bar{b})$	26.206 ± 0.028	1.0 ± 0.1
$\sigma(Z \to (\tau\tau) + b\bar{b})$	3.3 ± 0.1	0.13 ± 0.02

transverse momentum of the highest outgoing parton, and cross sections are quoted in $\hat{p}_T$ bins. Even at large values of $\hat{p}_T$ ($30 \div 50\,\mathrm{GeV}/c$) the cross sections for these processes remain high. If the b-jet identification fails and one (or more) leptons are picked up from the b-jets, all the signal final state objects are there.

The $Z(\to ll) + b\bar{b}$ final state is originated by an initial state that is an admixture of $gg(89\%)$, $u\bar{u}(7.7\%)$ and $d\bar{d}(3.2\%)$. The lowest order diagram, including all possible combinations of quarks in the initial state, a $b\bar{b} + e^+e^-/\mu^+\mu^-$ and both on- and off-shell Z, can be modeled using the COMPHEP generator [131], then exploiting PYTHIA [90] for showering and hadronization and to apply the basic kinematic cuts. The NLO k-factor is estimated with the Monte Carlo program MCFM [124], which makes full predictions for any infra-red safe variable, including fully differential distributions, for any set of experimental cuts. It turns out to be $k = 2.4 \pm 0.3$ when $p_T(l) > 10\,\mathrm{GeV}/c$, $|\eta(e)| < 5.5$. Selecting only $e^\pm$ and $\mu^\pm$ in the final state, the values in Tab. 3.1 are found [125]. The small contribution from the tau channel is due to the $\tau \to e/\mu\nu$ decays.

Since this process will be remarkably annoying only in the reconstruction of the Z mass invariant plot, only $M(l^+l^-) > 60\,\mathrm{GeV}/c^2$ will be addressed here. The $W(\to l) + b\bar{b}$ would need two fake leptons, and very few events may pass the Z invariant mass requirements, so this contribution is no more considered;

- *Z plus jets production*: This background source includes processes such as $q\bar{q} \to Zg$ and $qg \to Zq$ whose diagram are readable from Fig. 3.3(center and right) by replacing the qWq' with a qZq vertex. If some hard leptons are picked-up somewhere, the final state topology is prone to pass the qZ or bW selection requirements, hence a good b-jet identification and invariant mass constraints are mandatory. The cross section both for Z plus jets and W plus jets increases dramatically with decreasing $\hat{p}_T$ of the outgoing parton [126]. Table 3.2 from an ALPGEN [127] calculation gives the results in four $\hat{p}_T$ bins for the total Z+1 jet production [126]. The $Z \to \tau\tau$ decay has been generated but here it is considered inclusively, directly quoting its contribution to the e/μ rate. The tiny leptonic cross section in the last bin has no relevance, while the analysis itself

Table 3.2: The NLO cross-section values for the $pp \to Z(\to l^+l^-) + 1$ jet production, in four among the $\hat{p}_T$ bins relevant for the analysis. Theoretical error affects the last significant digit.

$\hat{p}_T$ range (GeV/c)	σ_{NLO}^{TOTAL} (pb)	$BR \cdot \sigma_{NLO}^{TOTAL}$ (pb)
50< $\hat{p}_T$ <85	983.7	97.39
85< $\hat{p}_T$ <150	304.8	30.17
150< $\hat{p}_T$ <250	69	6.83
250< $\hat{p}_T$ <400	3.7	0.37

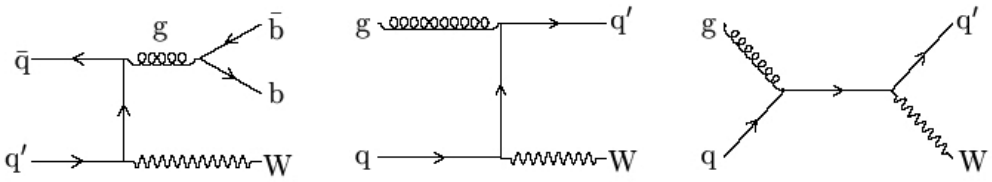

Figure 3.3: The lowest order diagram for the $pp \to W + b\bar{b}$ (left) and $pp \to W$+jet (center and right). Production of $Z + b\bar{b}$ and Z+jets occurs with similar diagragms, once charged vertexes have been replaced by neutral ones. Cross-section of this process strongly diverges with increasing $\hat{p}_T$.

will shown that the $\hat{p}_T$ <50 GeV/c bins are excluded from any selection. The other bins correspond to three different samples, in the analysis discussion indicated as $Z + j_{50-85}$, $Z + j_{85-150}$, $Z + j_{150-250}$. The Z+2 jets has in principle twice the chances to be mistagged with a genuine b-jets, but its cross section times branching ratio is about 2.8 pb for $\hat{p}_T$ >100 GeV/c. When surviving events from the cascade of cuts will be examined, it will turn out that the Z+1 jet is totally negligible. Therefore, the small Z+2 jets process can be safely ignored;

- *Di-boson production*: Events in which either a ZZ, ZW or WW pair is produced are natural candidates to pass the selection for a W and Z, mostly if they decay leptonically. Two LO diagrams for the $q\bar{q}' \to WW$, followed by a leptonic decay, are displayed in Fig. 3.4. The $q\bar{q}' \to ZW$ and $q\bar{q}' \to ZZ$ diagrams are obtained by replacing one or two of the qWq' with a qZq vertex, in only the t-channel for ZZ and in both t- and s-channel for ZW. However, the additional requirement of the presence of a b-jet strongly suppresses the contribution from this background source. The ZZ production has no large missing energy, thus it receives an additional suppression. Decays in τ leptons are another source of jets, leptons and missing energy so it is important to include them in the generation. Cross-sections for the three different combinations are listed below and milder cuts are considered here for lepton generation, with $p_T(e) > 5$ GeV/c, $|\eta| < 2.7$ and $p_T(e) > 3$ GeV/c, $|\eta| < 2.4$.

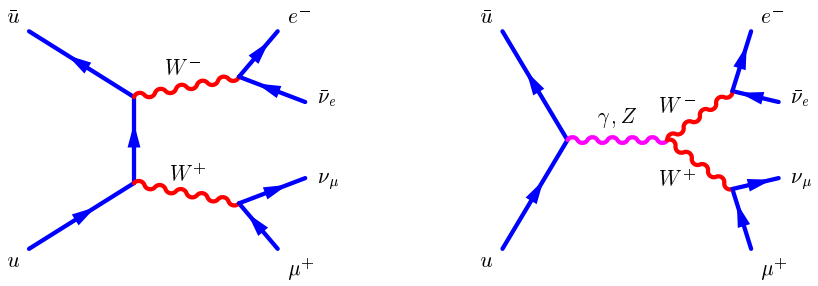

Figure 3.4: The lowest order diagrams for the $q\bar{q} \to WW$ with leptonic decays. The ones for $q\bar{q} \to ZZ/W$ can be obtained by replacing the qWq' with a qZq vertex.

Table 3.3: The LO and NLO cross-section values for the $pp \to ZW$ production, followed by Z and W leptonic decays. CTEQ5 parton distribution function is assumed.

	σ_{LO}^{TOTAL} (pb)	σ_{LO}^{TOTAL} (pb)	$BR \cdot \sigma_{NLO}^{TOTAL}$ (pb)
W^+Z	19.9	34.0	1.034
W^-Z	12.5	21.4	0.63

Cross sections from the three possible productions are:

- $pp \to WW + X \to 2l + X$:
  WW total cross section is 127.8 pb at NLO [128] when the CTEQ5 PDF is adopted. This results in a factor $k = 1.5$ greater than the LO value and receives contributions from $q\bar{q} \to WW$ (about 95%) and $gg \to WW$ (5%) [95]. Leptonic cross section is therefore 127.8 pb $\times BR(W \to l\nu)^2 = 13.4$ pb ($l = e, \mu, \tau$). The same process featuring 1/2/3 jets has a rate lower by a factor 1.4/2.1/3.7 (from ALPGEN): as the process with no jets itself will be proven to be totally negligible, these ones are not included in the study;
- $pp \to WZ \to 3l$:
  WZ production in p-p collisions proceeds mainly through s-channel quark annihilation. Cross-section for W^+Z and W^-Z are different. When CTEQ5 is adopted, LO and corresponding NLO cross section in Tab. 3.3 are delivered [128];
- $pp \to ZZ \to 4l$:
  The $pp \to ZZ$ production proceeds through the t-channel, with a $q\bar{q} \to ZZ$ dominant diagram. The total cross section for the ZZ production is 12.9 pb at LO and 17.2 pb at NLO [128]. For the purely leptonic channel, when only $q\bar{q} \to ZZ$ is considered one obtains $\sigma(ZZ \to 4l) = 0.168$ pb.
  The $gg \to ZZ$ gluon fusion contribution to the ZZ cross section is on average 10% of the LO quark-antiquark annihilation term [129], depending on the of the $4l$ final state. This is a small contribution on top of a not very relevant background

Table 3.4: The total and selected cross sections (all quoted in pb) for the processes considered relevant in the $t\bar{t} \to (Zq)(W\bar{b})$ decay channel. The symbol l denotes the sum of the μ, e and τ contributions. The last two columns quote the expected rate after $10\,\text{fb}^{-1}$ and the number of events generated in the analysis.

	σ^{TOT}(pb)	Notes	$\sigma^{NLO} \times BR$(pb)	Number of evs. ($10\,\text{fb}^{-1}$)	Number of generated evs.
$t\bar{t} \to 2l$	825	di-leptonic channel	88.4	884 k	908 k (TOPREX)
$Z + b\bar{b}$	794	$M(ll) > 60\,\text{GeV}/c^2$ $Z \to ll \to 2e/2\mu + X$	55.7	557 k	320 k (COMPHEP)
$Z + 1$ jet	1361	$40 < \hat{p}_T < 300\,\text{GeV}/c$, $Z \to ll$	134.8	1348 k	352 k (PYTHIA)
$WW + X$	127.8	di-leptonic channel, no hard jets	13.4	134 k	93 k (PYTHIA)
$ZW + X$	53.4	di-leptonic channel, no hard jets	1.66	16.6 k	49 k (PYTHIA)
$ZZ + X$	17.2	di-leptonic channel, no hard jets	0.168	1.68 k	93 k (PYTHIA)

source, thus it will be not considered in the generation procedure;

- *Single top production, $t \to l + X$*: Single top events, in which the top quark is produced in association with a W boson, a b quark or a forward light quark, could be prone to satisfy the requirements for the FCNC side (in addition to the SM ones). On the other hand, the requirement of three hard isolated leptons will heavily suppress this source of background, so only the dominant single-top process ($pp \to t + q + X$, t-channel) could have some relevance;

- *multi-jet production*: In the high-multiplicity LHC environment, QCD is one of the most challenging background for many analyses. Many jets initiated by light quarks (light-quark jets) will be produced, but the requirement of three hard isolated leptons is again powerful in suppression of this source. Therefore, QCD events will be not included in the $t \to qZ$ analysis.

Table 3.4 summaryzes all the background processes included in the analysis, with the cross sections quoted above, details about the specific generation channel and the number of events expected when the integrated luminosity reaches the reference value $10\,\text{fb}^{-1}$. The last column anticipates the number of events produced in the analysis by a specifc Monte Carlo generator.

Backgrounds for the $t \to \gamma q$ channel

The FCNC decay of top into a $(\gamma q) + (Wb)$ shares the same signatures with $(Zq) + (Wb)$ for what concern the missing energy from W and at least two hard jets – a b-tagged one from the SM top decay and a light one from the FCNC top decay. Instead of three hard leptons, the process has a single lepton and a single high energy photon, both isolated. Since the signal to be discovered takes no more advantage by the multi-lepton signature, a richer set of background sources have to be considered to reproduce a realistic situation. In general they can be searched among the following:

- the ones producing one or more hard leptons in association with one or more hard photons. When these objects are accompanied by some high energy jets, a signature similar to the signal could be obtained;
- the ones producing two or more b quarks and light hard jets, together at least with two hard leptons – one of which could be wrongly reconstructed as a photon.

It is worth to note that the photon signature may be produced by some hadron decay (mostly π^0) or mis-identified particles in jets, for which isolation cuts are essential.

The sources of background considered for this decay channel are listed below and their cross sections are reported in Table. 3.9.

- *$t\bar{t}$ production*: In this case the top quark which decays via the SM decay channel will be identified correctly, while the other top quark can mimic the FCNC top decay if an electron (from W or even from a b decay) is misidentified as the photon, or if a photon is produced in the decay of a π^0 close to the primary vertex. This effects implies that the whole $t\bar{t}$ production has to be included, because the photon radiation is independent from top decay and the misidentification with the signal may occur in all the top final states. The large $t\bar{t}$ production cross section makes the channel the most important source of background for this analysis;
- *Z and W plus jets production*: In this case, both Z+jets and W+jets may challenge the signal reconstruction. As in the $t \to Zq$ channel, these processes can easily mimic the signal, particularly if the b-tagging procedure fails. The cross section for W+jets is even larger that Z+jets but if only the relevant $\hat{p}_T$ bins are put in consideration, this background can be quite well controlled.

  Even for this process, the cross-section values can be quoted for different ranges in the transverse momentum of the recoiling parton. Table. 3.5 presents the NLO cross sections in six $\hat{p}_T$ bins, as estimated by the use of the MRST PDF [130]. The amplitudes for the the possible diagrams $q'\bar{q}/qg \to Wg/q$ are not significantly different and the cross section steeply decreases with p_T.

Table 3.5: The NLO cross-section values for the $pp \to W + 1$ jet production, in four among the $\hat{p}_T$ bins relevant for the analysis. Theoretical error affects the last significant digit.

$\hat{p}_T$ range (GeV/c)	σ_{NLO}^{TOTAL} (nb)	$BR \cdot \sigma_{NLO}^{TOTAL}$ (pb)
0< $\hat{p}_T$ <20	313.7	109800
20< $\hat{p}_T$ <50	31.9	11165
50< $\hat{p}_T$ <85	6.93	2426
85< $\hat{p}_T$ <150	2.20	770
150< $\hat{p}_T$ <250	0.37	129.5
250< $\hat{p}_T$ <400	0.0542	19.51

The theoretical error ranges from about 4.3% to 1.4% for the highest energy bins. If same parton $\hat{p}_T$ are compared with the Z+jets, the W+jets has a rate greater than one order of magnitude: therefore, severe cuts at the selection level are mandatory. The more efficient way is to tightly constrain the jets to have a high probability to come from a b quark. As the bottom production is much suppressed in these boson+jet events, such requirements lower the rate to a manageable level. Once these operations have been performed, the production of Z/W plus more than one jet becomes negligible;

- *Vectorial boson production in association with a $b\bar{b}$ pair*: Together with the $Z + b\bar{b}$ processes mentioned above, the $W + b\bar{b}$ displayed in Fig. 3.3(left) may be of concern here, because one of two b-jets could not be tagged and a fake photon has a chance to be found in the multi-jet production.

  Next-to-leading order calculation are particularly important in this case, because they introduce initial interactions different from the $q\bar{q}' \to Wg$ (via gluon radiation), and strongly reduce the cross-section dependence from the W transverse momentum. The inclusive cross sections, evaluated both at LO and NLO with MCFM, are in Tab. 3.6, where the uncertainty is dictated by the factorization scale varied between $M_W/2$ and $2M_W$. The remarkable scale differences between LO and NLO are well shown. Results obtained with NLO generators as ALPGEN and MADGRAPH that may operate on exclusive states with n jets are in good agreement with these value and between them;

- *Di-boson production*: Events in which a ZZ, ZW or WW pair is produced are an important source of hard leptons even in this case. High energy photon could be produced by bremsstrahlung of a high p_T electron, or be signaled by a mis-identification in the photon reconstruction algorithm. Since these effects will occur few times, the additional request for a b-jet will completely suppress this component. Therefore, even in this case the only di-boson di-leptonic channel are considered relevant;

Table 3.6: The inclusive cross sections for the $W + b\bar{b}$ production resulting in electron and missing energy. They are shown at LO and NLO and the error induced from the factorization scale. Leptonic universality is assumed, so an equal rate for muon production is expected.

	σ_{LO} (pb)	σ_{NLO} (pb)
$W^-(\to e^-\bar{\nu}) + b\bar{b}$	$1.30^{+0.21}_{-0.18}$	$3.06^{+0.62}_{-0.54}$
$W^+(\to e^+\nu) + b\bar{b}$	$0.90^{+0.14}_{-0.11}$	$2.11^{+0.46}_{-0.37}$

Table 3.7: Cross-sections for the $W/Z + \gamma$ production at LO, NLO and scaled for the observed channel branching ratio. Though the final state may emulate the signal, the tiny resulting rate makes the process totally negligible.

	σ_{LO}	σ_{NLO}	$BR \cdot \sigma_{NLO}$
$W^- + \gamma$	3.08 pb	10.15 pb	2.23 pb
$W^+ + \gamma$	4.79 pb	13.89 pb	3.06 pb
$Z^0 + \gamma$	1.29 pb	2.37 pb	0.16 pb

- *Single top production, $t \to l + X$*: As for the $t \to qZ$ channel, single top events may be a source of background if the particles produced in association with the top quark could be misidentified as a $t \to q\gamma$ decay. In the present channel, this is much more likely because it is not protected by the multi-lepton request, and the additional photon can be faked by final state particles accompanying the top, like light-jets and leptonic W decays. The processes that could have a relevance here are the t- and the tW-channel with the top and W decay in e, μ, resulting in $\sigma_{EW}(t\text{-}ch.) \times BR = 246 \times 0.35 = 86.1$ pb and $\sigma_{EW}(tW\text{-}ch.) \times BR = 64 \times (0.35)^2 = 7.84$ pb. The s-channel single-top, beside a tiny cross section $\sigma_{EW}(s\text{-}ch.) \times BR = 10.6 \times 0.22 = 3.7$ pb, features two jets coming from b quark, so it is expected to be suppressed by a multi-b-jet veto;

- *Boson plus photon production*: The production of events containing either a W or a Z in association with a hard isolated photon should also be considered. The prediction for cross section is quoted in Tab. 3.7, where the uncertainty is on the last digit and the photon is considered to have $p_T > 50\,\text{GeV}/c$ (for the $W + \gamma$) and $p_T > 100\,\text{GeV}/c$ (for the $Z + \gamma$). These cross sections rapidly decrease with the photon transverse energy. As in the analysis below a $p_T > 80 \div 100\,\text{GeV}/c$ is required, the rates of the processes are extremely low and should be completely rejected by the requirement of one b-jet per event. On this basis, this background contribution is totally neglected;

- *Multi-jet production*: QCD background has a much more severe impact in this case

Table 3.8: Cross-sections (in pb) for the multi-jet production in the whole kinematic range from 50 to 1000 GeV/c, partitioned in ranges of $\hat{p}_T$. Also this process exhibits a dramatic increasing in rate when $\hat{p}_T$ is lowered. The degree of suppression of such background source will be discussed separately in this chapter. The last column indicates the number of events generated with PYTHIA.

$\hat{p}_T$ range (GeV/c)	σ_{NLO}^{TOTAL}(pb)	Number of generated evs.
QCD (50< $\hat{p}_T$ <80) GeV/c	20,917,810	92 k
QCD (80< $\hat{p}_T$ <120) GeV/c	2,946,763	96 k
QCD (120< $\hat{p}_T$ <170) GeV/c	499,156	93 k
QCD (170< $\hat{p}_T$ <230) GeV/c	100,800	93 k
QCD (230< $\hat{p}_T$ <300) GeV/c	24,470	93 k
QCD (300< $\hat{p}_T$ <380) GeV/c	6,384	93 k
QCD (380< $\hat{p}_T$ <470) GeV/c	1,887	96 k
QCD (470< $\hat{p}_T$ <600) GeV/c	683	96 k
QCD (600< $\hat{p}_T$ <800) GeV/c	204	90 k
QCD (800< $\hat{p}_T$ <1000) GeV/c	35.1	80 k

compared to the $t \to Zq$ channel, because a large number of photons are normally produced inside jets. To judge whether these kind of processes are sizable or not a special care is needed, that requires to make use of the selection efficiencies determined by the analysis. This *a posteriori* evaluation will be done at the end of the chapter, where reduction factors found with simulated events will be applied to QCD and other possible contaminations.

The cross section for the multi-jet production decreases rapidly with the $\hat{p}_T$ of the outgoing parton, which means that the sensitivity of the analysis to such background source depends on the hard scale. Therefore, a number of events samples fully covering the range $50 < \hat{p}_T < 1000$ GeV/c needs to be studied. Cross-sections in several bins are reported in Tab. 3.8, as the total NLO value expressed in pb. A small part of these kind of events (quoted in the last column) has been generated and constitutes a subsample that will be discussed separately.

Table 3.9 summarizes the background processes included for the FCNC signal with γ. The cross sections, details about the specific generation channel and the number of events expected when the integrated luminosity reaches the reference value 10 fb^{-1} are reported, as well as the rate simulated with the Monte Carlo generators.

The processes introduced here and listed in Tab. 3.4 and Tab. 3.9 have been generated with a size as close as possible to the rates at 10 fb^{-1}. Some details on such a step are sketched below.

Table 3.9: The total and selected cross sections (all quoted in pb) for the processes considered relevant in the $t\bar{t} \to (\gamma q)(W\bar{b})$ decay channel. The symbol l denotes the sum of the μ, e and τ contributions. The last two columns quote the expected rate after $10\,\mathrm{fb}^{-1}$ and the number of events generated in the analysis.

	σ^{TOT}(pb)	Notes	$\sigma^{NLO} \times BR$(pb)	Number of evs. ($10\,\mathrm{fb}^{-1}$)	Number of generated evs.
$t\bar{t} \to 2l$	825	inclusive process	825	8250 k	2500 k (TOPREX)
$Z + b\bar{b}$	794	$M(ll) > 60\,\mathrm{GeV}/c^2$, $Z \to ll \to 2e/2\mu + X$	55.7	557 k	320 k (COMPHEP)
$Z + 1$ jet	1361	$40 < \hat{p}_T < 300\,\mathrm{GeV}$, $Z \to ll$	134.8	1348 k	352 k (PYTHIA)
$WW + X$	127.8	di-leptonic channel, no hard jets	13.4	134 k	93 k (PYTHIA)
$ZW + X$	53.4	di-leptonic channel, no hard jets	1.66	16.6 k	49 k (PYTHIA)
$ZZ + X$	17.2	di-leptonic channel, no hard jets	0.168	1.68 k	93 k (PYTHIA)
$W + b\bar{b}$	–	$W \to e/\mu + \nu$ no hard jets	5.17	51.7 k	50 k (PYTHIA)
$W + 1$ jet	2200	$80 < \hat{p}_T < 150\,\mathrm{GeV}/c$, leptonic channel	770	7700 k	3720 k (PYTHIA)
single-top	246 64	t-channel, $W \to l\nu$ tW-channel, $W \to l\nu$	86.1 7.84	861 k 78.4 k	384 k 49 k (TOPREX)

3.1.3 Details of signal and background generation

All the hard processes outlined above are generated with PYTHIA 6.324, TOPREX 4.11 [91] and COMPHEP . They evaluate the cross section values as well as the incoming and outgoing particles momenta and colours. Subsequently, these events are evolved through a parton showering and hadronization step and the decay of the unstable particles, that is is performed by PYTHIA .

All samples are generated assuming pp collisions with $\sqrt{s} = 14$ TeV and the following input values for the physical variables:

$$\begin{aligned} M_Z &= 91.1876\,\text{GeV}/c^2 & \Gamma_Z &= 2.49\,\text{GeV}/c^2 \\ M_W &= 80.425\,\text{GeV}/c^2 & \Gamma_W &= 2.06\,\text{GeV}/c^2 \\ BR(Z \to ll) &= 0.099 & BR(W \to l\nu) &= 0.350 \\ m_t &= 175\,\text{GeV}/c^2 & \Gamma(t \to bW)/V_{tb} &= 1.55\,\text{GeV} \\ m_b &= 4.8\,\text{GeV}/c^2 & m_c &= 1.27\,\text{GeV}/c^2 \\ \alpha_{EM}^{-1} &= 128.89 & \alpha_S(M_Z) &= 0.118\ . \end{aligned} \tag{3.1}$$

The default PDF is the CTEQ5L, in which running of α_S is defined at first order; structure functions are used with the evolution parameter Q^2. Q^2 is equal to $m_A^2 + p_T(A)^2$ with $A = t$ for the signal and $A = W/Z$ for the background with a gauge boson and a t-channel single-top, $Q^2 = \hat{p}_T$ for the QCD background. In all samples, electrons are generated down to $p_T > 5\,\text{GeV}/c$, $|\eta| < 2.7$ and muon down to $p_T > 3\,\text{GeV}/c$, $|\eta| < 2.4$.

Production with TOPREX

The FCNC signal samples were produced using the specialized event generator TOPREX [91].

TOPREX is a generator providing (LO) matrix elements for several hard processes not implemented in PYTHIA , mainly involving top quark. Top spin polarizations are taken into account and a correct description of the differential distributions and correlations of the top quarks decay products is implemented. Most of top quark decay channels (observed or not), including gauge bosons and charged Higgs are fully supported.

In TOPREX routines, all anomalous couplings are set equal to one, *i.e.* $|v_{tq}^Z|^2 + |a_{tq}^Z|^2 = \kappa_{tq}^Z = \kappa_{tq}^\gamma = 1$, and $v_{tq}^Z = a_{tq}^Z = 1/\sqrt{2}$ (see Sec. 1.3). This results in a total width $\Gamma = 1.437\,\text{GeV}$ and a branching ratio $BR(t \to Z/\gamma q) = 0.325$, accordingly with results in Tab. 1.4.

The generated signal sample is an admixture of the two charge-conjugated final states $t\bar{t} \to (Vq)(W\bar{b})$ and $t\bar{t} \to (Wb)(V\bar{q})$ (in equal part) and amounts to 8000 events each for the $V = Z^0$ and $V = \gamma$ channels.

Production with CompHEP

CompHEP is a package for evaluating Feynman diagrams for SM and several versions of MSSM. It allows for the computation of scattering processes with up to 6 particles and decay processes with up to 7 particles in the final state. In the present analyses, it has been adopted only for the generation of the $Zb\bar{b}$ process.

Interface to Pythia

The Pythia package is a general-purpose generator for hadronic events in pp, e^+e^- and ep colliders. It contains a sub-processes library and generation machinery, initial- and final-state parton showers, underlying event, hadronization and decays. All subprocesses are available at leading order. Pythia accesses to the hard processes (externally or internally generated) and takes care of parton showering, hadronization, fragmentation and decays of the unstable particles.

All events have been generated using default parameters for the hard process and subsequent fragmentation. Here some of the more relevant ones are listed:

Parton Showering: Initial and final state QED and QCD (ISR, FSR) radiation is evolved up to a maximum parton virtuality Q^2 in in space-like showers, or up to m^2 when a resonance with mass m is present;

Hadronization: The set of colour singlet quark-antiquark pairs are merged into hadrons via a the 'Lund string model' [132] for the light flavours, while the Peterson/SLAC function [133] is used for charm and heavier ones;

Final state: Decays are switched on only for unstable particles with a decay length for which $c\tau < 10\,\text{mm}$, that is those containing a charm or bottom quark. Thus decays of hadrons as $K^0_S, \Lambda^0, \Sigma^{\pm,0}, \Xi^{-,0}, \Omega^-$ are not generated by Pythia but managed by the CMS detector simulation. Final state resonances (with mass greater than 20 MeV) are smeared by a Breit-Wigner distribution;

Multiple interactions: The multiple parton interactions model assumes a hadronic matter overlap consistent with a double Gaussian distribution. A 0.4 fraction of the total radius is supposed to be contained in the 0.5 fraction of total hadronic matter. The effective transverse momentum for multiple interactions p_T scales as $(p_T/1\,\text{TeV})^{0.16}$, starting from a minimum of $1.9\,\text{GeV}/c$. The probability that an additional interaction in the multiple interaction formalism gives two gluons (with colour connections to nearest neighbours in momentum space or as a closed gluon loop) is 0.66. Remaining fraction is supposed to consist of quark-antiquark pairs.

As already specified, the well-known limit of Monte Carlo parton-showers generator as Pythia is the approximation of matrix elements at the leading order. The simple introduction

of k-factors is not always the solution, because the inclusion of higher orders may also involve distortions in differential distributions. In the present analyses, theoretical errors on the cross sections – that in principle cannot be evaluated unless the higher order calculation is available – is not a major issue, since the amount of background will be directly evaluated from data. Possible modifications to differential distribution will be addressed in the next chapter, where it will be demonstrated that the job performed by PYTHIA is good enough for the scope.

3.2 The physics objects reconstruction

This section comes before the specific descriptions with details of the analyses, because it contains most of what is common to the two top decays studies. Therefore, here below there is a description of the algorithms adopted to reconstruct the 'physics objects' involved in both the addressed channels, *i.e* electrons, muons and jets, and consequently missing transverse energy and jets with a beauty flavour. In most of case, these algorithms are developed using variables provided by the reconstruction code. For cuts optimization purpose, comparing the reconstructed objects with the 'true' ones from simulated event samples information is always very useful. Namely, founding a reconstructed final state object (a lepton, a jet, or a missing transverse energy vector) very close to the true particle is a demonstration that the reconstruction algorithm performs at its best. Where no differently explicited, reconstructed objects are intended to be 'matched' when the distance in the $\phi - \eta$ space is $\Delta R < 0.3$.

3.2.1 Lepton reconstruction and identification

Both the analyses addressed by this work require an efficient identification of muons and electrons. The reconstruction of a final state with three leptons (as in $t\bar{t} \to W(\to l\nu)Z(\to ll)+X$) or with a lepton and a hard photon (as in $t\bar{t} \to W(\to l\nu)\gamma+X$) in presence of a multi-lepton production background, could be a tricky task and requires a careful optimization of selection parameters.

The sample of e and μ coming from vector boson can be easily contaminated by leptons from different sources. Among them, the most important are:

- the prompt decay of a bottom hadron, that constitutes the main contribution of background when energetic leptons ($p_T > 20 \div 30\,\mathrm{GeV}/c$) are selected, and the prompt decay of a primary charm hadron;
- the prompt decay of a charm hadron, originating from a bottom hadron. Standard Model processes having one or more b-jet (as $Zb\bar{b}$ and $t\bar{t} \to WbW\bar{b}$, single-top decay) will be the more important issues;
- the decay of light hadrons coming either from the primary hadronization or from the decay of heavier hadrons.

Isolation is the main tool to get rid of this contamination. As it will be explained, in each analysis muons and electrons isolation (along with other variables) is optimized, in order to keep the background at a minimum without resulting in a significant decrease of the efficiency.

Electron reconstruction and identification

Electrons to be pre-selected are those reconstructed from the algorithm sketched in Sec 2.3.1. These objects still retain a too big contamination with the mentioned background processes – especially electrons from heavy flavour decays and from photoconversions. Further quality cuts that can enforce the "electronicity" of the candidate electrons are demanded. They are briefly explained in the following.

In order to fully exploit the information from both the tracker and ECAL, it is implicit that $|\eta(e)| < 2.6$ is always assume. This is the region where the two detectors can be matched. All tracks used in the analyses have been reconstructed using a combination of the pixel detector and the silicon strip tracker; transverse momenta p_T are measured from them.

Electron isolation:
The simplest and most powerful isolation criteria is obtained from tracks originating from a common (primary) vertex. In the whole studies, tracks are required to have $|\Delta z|$ <0.4 cm, where $|\Delta z| = |z(e) - z(tk)|$ is the difference between the longitudinal impact parameter and the z position of the primary vertex, and $|\Delta_{xy}|$ <0.1 cm, where $|\Delta_{xy}| = \sqrt{(x(e) - x(tk))^2 + (y(e) - y(tk))^2}$.

In order to not include 'ghost' tracks in the isolation, a good track is defined as one which satisfies $p_T > 0.9\,\mathrm{GeV}/c$ and has more than 4 hits in the tracking detectors associated to it. The isolation cone in the (η, ϕ) space is taken as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.1$: this tight cone can be assumed as an optimal choice for the best efficiency/background rejection when handling hard electrons (see for example Ref. [102], Sec. 10.4.7). The electron isolation variable (hereafter referred as $Isol(e)$) is then defined as the sum of the p_T of all the tracks satisfying these requirements and found inside the cone – excluding the electron track – divided by the electron p_T. In these analyses, avoiding to enter the electron track is obtained by excluding a $\Delta R > 0.01$ around it in the sum of transverse momenta.

Other electron isolation variables (*e.g.* based on cluster isolation) are not involved in any part of the work, the main focus being to identify the tracks around the electron one – mainly those from b quark decays.

Geometrical matching between tracks and superclusters:
The supercluster-driven pixel seed finding has generally a loose matching between the reconstructed electron track and the corresponding supercluster. On the other hand, the matching between the total energy E_{SC} collected by the supercluster with the momentum measured at the track origin is sensitive to the energy lost in the tracker material: as a consequence, it is worth to inquire the discrimination power of some matching variables. The choices adopted

here are:

- the η geometrical matching $|\Delta\eta| = |\eta(\text{SC}) - \eta(\text{extrap. track})|$, where $\eta(\text{SC})$ is the supercluster η position and $\eta(\text{extrap. track})$ is the track pseudorapidity at the closest position to the supercluster. It is demonstrated to efficiently reject the fake electron candidates found in QCD jets;
- the ϕ geometrical matching $|\Delta\phi| = |\phi(\text{SC}) - \phi(\text{extrap. track})|$, where $\phi(\text{SC})$ is the supercluster ϕ position and $\phi(\text{extrap. track})$ is the track azimuthal angle at the position closest to the supercluster.

When more than one cluster is found closer than $\Delta R = 0.03$, the one with the better $\Delta\phi$ and $\Delta\eta$ is assumed to be associated with the hard electron, and the others are discarded.

Energy matching between tracks and superclusters:
In order to select most of "golden electrons" (see Sec. 2.3.1) having more chances to originate from an heavy gauge boson, the corrected energy measurement E_{rec} provided by electromagnetic calorimeter can be combined with the tracker momentum measurement. This also improves the estimate of the electron momentum at the interaction vertex.

When considering electron candidates, the energy of the cluster in the calorimeter have to be equal within measurements uncertainties to the corresponding track momentum. As it will be demonstrated for these specific analyses, for the signal electron the bulk of E_{rec}/p_{in} distribution (where p_{in} is the transverse momentum measured with track parameters at the vertex) is roughly between 0.9 and 1.2.

Cases with $E_{rec}/p_{in} > 1$ are usually due to an underestimation of the track momentum. They may be originated by an early emission of a large amount of radiation along the electron trajectory, that weakens the track but is fully recovered by the supercluster. Cases with $E_{rec}/p_{in} < 1$ can be imputable either to an underestimate of the true energy by the calorimetry measurement (*e.g.* due to energy lost from the conversion of secondary photons) or to an overestimate of the true initial track momentum. A worse matching between E_{rec} and p_{in} indicates the candidate is likely not to be an electron from W/Z, for example there is some relevant hadronic activity around it. Optimized cuts will be presented later for the $t \to Zq$ and $t \to \gamma q$ analyses.

Energy matching between ECAL and HCAL:
The energy deposits from an electron will tend to be fully contained within the electromagnetic calorimeter, while low p_T hadrons ($p_T > 20 \div 30\,\text{GeV}/c$) deposit energy partially in the electromagnetic and (mostly) in the hadron calorimeter. The ratio of energy deposited in HCAL (in the region defined by the hadronic trigger tower behind the super-cluster crystal with highest energy, E_{HCAL}) to that deposited in the ECAL (E_{ECAL}) may be useful to tag multi-jet activity accompanying the electrons. As this activity often originates from back-

ground (*e.g.*, neutral hadrons), an upper limit on E_{HCAL}/E_{ECAL} is another good handle for the electron pre-selection. 'Genuine' electrons usually do not exceed $E_{HCAL}/E_{ECAL} \sim 0.05$.

Muon reconstruction and identification

In general, the selection of hard muons for the off-line analysis is rather simple in CMS – that is a tool well-conceived to identify muons. The redundancy of the muon system, along with the robust matching procedure with the tracker, allows the objects delivered by the muon trigger to be mostly 'real' muons. As a result, a pre-selection based on a simple p_T threshold is usually enough to reach a very high reconstruction efficiency.

The off-line muon algorithm explained in Sec. 2.3.2 comes in help in several case and is included in the reconstruction code. Low-p_T muons, for example, range out within the iron yoke before depositing hits in the outer muon detector layers, so they are difficult to reconstruct with the standalone muon track reconstruction. This class of muons can potentially be identified offline via this algorithm, by matching the hits found in the inner muon detector layers with reconstructed silicon tracks, or examining the associated calorimeter energy deposits to see if they are compatible with those from a minimum-ionizing particle.

In these analyses, all muons needed for the final state reconstruction have to be isolated. The isolation cone drawn around the direction of a pre-selected muon may be quite larger than the one for electron: several optimization studies (see for instance Ref. [102], Sec. 9.3.2) dictated about $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$, that has shown to be robust even in the high luminosity phase. Similarly to the electron case, tracks are included in the cone only if they have $p_T > 0.9\,\mathrm{GeV}/c$ and more than 4 hits in the tracker. As it will be measured on simulated reconstructed muons, the direction resolution is quite good for both algorithms ($\sigma(\phi) \approx \sigma(\eta) \approx 10^{-3}$), so a very small veto cone can be choose. The same veto cone used for the electron $\Delta R = 0.01$ demonstrates to perform very well. To attain the strongest isolation requirements, no matching in transverse plane and in longitudinal direction between the lepton track and the tracks inside the cone is applied. This allows the inclusion in the cone also of objects from semileptonic b quark decays, that are expected to be detached from the primary vertex.

3.2.2 Jets reconstruction and identification

All jets are reconstructed using an iterative cone algorithm with a radius of $\Delta R < 0.5$, which is ran over transverse energy deposits in the CMS calorimeters ($\vec{E}_T(tower)$). The "raw" jet energies determined in this way ($E_T^{raw}(jet)$) are then calibrated to the jet energies $E_T^{calib}(jet)$ through the use of the "gamma-jet" method [114]. Such calibration issues a system of expression to rescale E_T^{raw} to E_T^{calib}, that holds in three different energy regions:

Table 3.10: The calibration constants for the rescaling of the energy in the "raw" jet. The calibration results come from the "gamma+jet" method.

	E_1 (GeV)	E_2 (GeV)	h_A (GeV $^{-1/2}$)	E_A (GeV)	k_A	h_B (GeV $^{1/2}$)	k_B	E_B (GeV)
$\|\eta\| < 0.226$	60	80	0.0554	37.24	0.0970	-3.522	0.853	-1.174
$0.226 < \|\eta\| < 0.441$	90	120	0.0400	6.743	0.3018	-2.254	0.100	7.044
$0.441 < \|\eta\| < 0.751$	70	90	0.0576	59.79	-0.0027	-2.791	0.731	17.534
$0.751 < \|\eta\| < 0.991$	50	70	0.0414	3.088	0.3013	-2.723	0.683	23.851
$0.991 < \|\eta\| < 1.260$	100	120	0.0372	12.92	0.2666	-2.445	0.919	-22.904
$1.260 < \|\eta\| < 1.496$	75	85	0.0512	15.73	0.1771	-3.719	2.136	-50.000
$1.496 < \|\eta\| < 1.757$	60	80	0.0762	46.37	0.0109	-3.880	2.931	-36.522
$1.757 < \|\eta\| < 2.046$	60	80	0.0555	12.09	0.2736	-2.821	2.113	-4.678
$2.046 < \|\eta\| < 2.487$	50	60	0.0353	-6.462	0.5447	-1.682	1.567	-10.685
$2.487 < \|\eta\| < 2.690$	45	55	0.0183	-3.342	0.6965	-22.62	494.2	-1.000
$2.690 < \|\eta\| < 2.916$	35	45	0.0840	2361	-3.292	-22.74	1518	-1.000

$$
\begin{aligned}
E_T^{calib}(E_T^{raw}) &= \frac{E_T^{raw}}{h_A\sqrt{E_T^{raw}+E_A}+k_A} \qquad \text{if } E_T^{raw} < E_1, \\
E_T^{calib}(E_T^{raw}) &= \frac{E_X}{E_T^{raw}+E_C}E_T^{raw} \quad \text{if } E_1 < E_T^{raw} < E_2, \\
E_T^{calib}(E_T^{raw}) &= \frac{E_T^{raw}}{\frac{h_B}{\sqrt{k_B E_T^{raw}+E_B}}+1} \qquad \text{if } E_T^{raw} > E_2.
\end{aligned}
\tag{3.2}
$$

All the parameters depend from the pseudorapidity of the jet. They can be quoted in bins of $|\eta|$ and their values are reported in Tab. 3.10.

A last parameter E_C is instead a function of the others, as:

$$
E_C = \frac{k_C E_2 - E_1}{1-k_C}, \quad \text{where} \quad k_C = \frac{h_A\sqrt{E_1+E_A}+k_A}{\frac{h_B}{\sqrt{k_B E_2+E_B}}+1} \quad .
$$

The E_X parameter is also a rather complicated function of the others, and it is not indicated here.

Missing transverse energy reconstruction

General concepts concerning E_T^{miss} definition and uncertainty sources have been sketched in the previous chapter. Equation 3.3 is the baseline recipe to evaluate missing energy, and it is customary to rewrite it as [115]:

$$
\vec{E_T^{miss}} = -\left[\vec{E}_T^{raw}(jet) + \sum_{jets}\left(\vec{E}_T^{corr}(jet) - \vec{E}_T^{raw}(jet)\right)\right] \quad , \tag{3.3}
$$

where it is evidenced that the full energy vector (*i.e.*, the x and y components) has to be known.

The reconstruction CMS framework offers few different choices for reconstructing E_T^{miss}, corresponding to different algorithms adopted and objects involved in its evaluation. Exploiting these algorithms, the missing transverse energy is obtained in several steps:

- the "ECAL + HCAL" method basically performs a vectorial sum of the uncalibrated (raw) towers in both the calorimeters. The opposite of this sum is taken as the missing transverse energy. In both the analyses addressed here, raw jets are only used if $\vec{E}_T^{raw} > 10\,\text{GeV}$;
- the expression in Eq. 3.3 is evaluated, adding the transverse energy of all the calibrated jets, and subtracting that of uncalibrated ones to the raw energy;
- in the "muon + jet" method, the previous estimate is corrected for the muon energy contribution;
- in the "MET + IC" method, the magnitude of the missing transverse energy is a simple vectorial sum of calibrated jets, clusterized with the IC algorithm.

In the two analyses the performance of these different definitions will be discussed and compared with the neutrino transverse momentum, that is the ultimate goal of this reconstruction.

In order to assess the use of the missing energy definition with better performances, here the E_T^{miss} variables are compared as they come from the calibrated towers method (CT), the calibrated towers with correction from jet energy method ($CTCorr$), the sum of jets as delivered from the IC method (IC) and the "muon + jet" correction method (MJ). In Fig. 3.5(left), all these four definitions are compared with the p_T of a generated neutrino up to $250\,\text{GeV}/c$.

Figure 3.5(right) shows the quality of the agreement by another point of view. The distribution of the difference between E_T^{miss} and $p_T(\nu)$ is represented for the same four definitions, and its mean and r.m.s. is an indication of the missing energy resolution. If no corrections to the towers vectorial sum or to the IC jet energy are applied, neutrino transverse energy is overestimated for jets having an energy exceeding 50 GeV, while in the softer part (where energy from calorimeter towers tend to be un-clustered) it results badly underestimated. This induces a cumulative energy bias in the whole range about 17 GeV for CT and 9 GeV for IC jets.

The correction on the calibrated jets recovers a large part of the unclustered energy, reaches a good agreement with clustered jets and the bias shifts back to about 5 GeV. The use of "muon + jet" method, though is not mandatory when $E_T^{miss} > 50\,\text{GeV}$, allows to model some low energy effects more precisely, and the average bias shrinks to 2 GeV. Therefore, the MJ is found to give the distribution closest to the $p_T(\nu)$ distribution of the neutrinos from

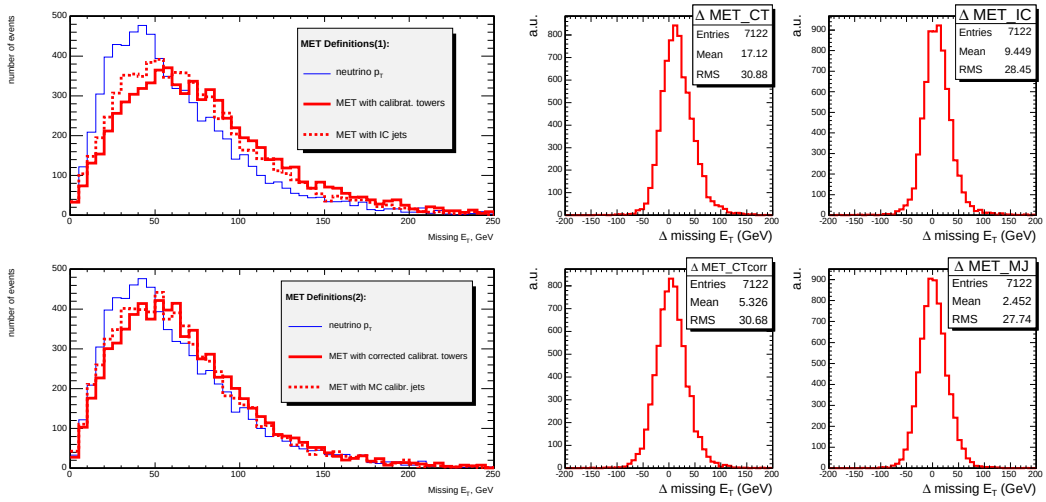

Figure 3.5: Left: A comparison between the transverse momentum of the generated neutrino with the missing transverse energy reconstructed by four different methods (provided by the reconstruction framework or developed in the analysis). Right: The distribution of the difference between the p_T of the generated neutrino and the missing energy, as calculated from the same four methods. The quality of the reconstruction is related to the width of these histograms.

the decay of the W, and it will be adopted in both the present analyses. The width of the obtained distributions results from the convolution of the missing energy distributions with its resolution, that is expected to follow a rule like $E_T^{miss} \approx 1.25\sqrt{\sum E_T}$.

When a b-jet is combined with the reconstructed W candidate, the mass of the top quark that produced it can be determined only if the longitudinal component of the missing energy vector is known. Since there are no means to measure the $p_z(\nu)$ like the $p_T(\nu)$ because is not possible to fully recover the longitudinal missing energy, kinematic constraints have to be required and this variable analytically determined. If one assumes that $E_T^{miss} = p_T(\nu)$ and neutrino, together with a lepton, comes from a W with a well known mass, $p_z(\nu)$ can be found by solving the equation $(E(l) + E(\nu))^2 - |\vec{p}(l) + \vec{p}(\nu)|^2 = M_W^2$, where $E(l/\nu)/\vec{p}(l/\nu)$ are the total energies/momenta of lepton and neutrino. As $\left(E_T^{miss}\right)^2 = p_x(\nu)^2 + p_y(\nu)^2$ and $p_T(l)^2 = p_x(l)^2 + p_y(l)^2$, the previous equality is solved by the following quadratic equation:

$$M_W^2 = 2\left(E(l)\sqrt{p_z^2(\nu) + (E_T^{miss})^2} - p_x(l)p_x(\nu) - p_y(l)p_y(\nu) - p_z(l)p_z(\nu)\right).$$

Since the equation is quadratic, there are two possible solutions:

$$p_z^{1,2}(\nu) = \frac{Ap_z(l) \pm \sqrt{\Delta}}{p_T(l)^2},$$

where

$$A = \frac{M_W^2}{2} + p_x(l)p_x(\nu) + p_y(l)p_y(\nu) \quad \text{and} \quad \Delta = E(l)^2 \left(A^2 - \left(E_T^{miss} \right)^2 p_T^2(l) \right).$$

The solution which gives the smallest value for $|p_z(\nu)|$ is assumed to be the correct one and is used to reconstruct the W candidate four-vector. The correctness of this option will be measured case by case with a comparison with true W boson generated from Monte Carlo.

When the Δ parameter evaluated with the variable from the reconstruction procedure becomes negative, the possibility to have no real solutions opens up. This un-physical situation is a consequence of limited E_T^{miss} resolution and the presence of "extra" missing energy from, for example, neutrinos from semi-leptonic heavy quark decays or particles which lie outside the detector acceptance. When such situation occurs (in about 30% of signal events), the choice to force $\Delta = 0$ is adopted. It can be considered as a reasonable option, because it corresponds to suppress the imaginary part of the solution, cleaning the reconstruction from these spurious effects.

3.2.3 b-tagging and anti-tagging of jets

Both the final states simulated in this work include the SM decay of one top quark, so every signal event is supposed to contain only one b-jet. On the other hand, the production of $t\bar{t}$ followed by two SM decays $t\bar{t} \to Wb + W\bar{b}$ (with an inclusive or di-leptonic final state) is one of the most annoying background, and is distinguished by the presence of two b-jets. This implies that a good b-jet identification is crucial for these analyses, both to suppress the non-top sources of background, and to help in reducing the contribution from the SM $t\bar{t}$ background.

In the following, the general strategy adopted in identifying the b-jets with the best purity and efficiency is outlined. The properties of the single b-jet in the signal have no differences in the two channels considered here, because FCNC possible decay channels do not significantly affect the kinematic of the SM side. On the other hand, relevant background is in principle different, dictating different choices in the selection parameters. Hence, details on the b-tagging cuts have to be treated separately for the two analyses, later in this chapter.

Selection of b-jets

In the ORCA simulated data, jets with a b flavour have been reconstructed using the procedure outlined in Sec. 2.3.4. The "combined b-tagging" algorithm is applied to calibrated jets reconstructed with the Iterative Cone method ($\Delta R < 0.5$), having a significant transverse momentum ($p_T > 20\,\mathrm{GeV}/c$) and $|\eta|$ <2.5, in order to exclude $b\bar{b}$ from gluon splitting and other QCD effects. Then a 'discriminator' parameters $Disc$, intended as the logarithm of the Likelihood Ratio between the b-jet and the non-b-jet hypothesis, is extracted for each of these. In addition, the number of tracks for each b-jet object is provided and it may constitutes an additional handle.

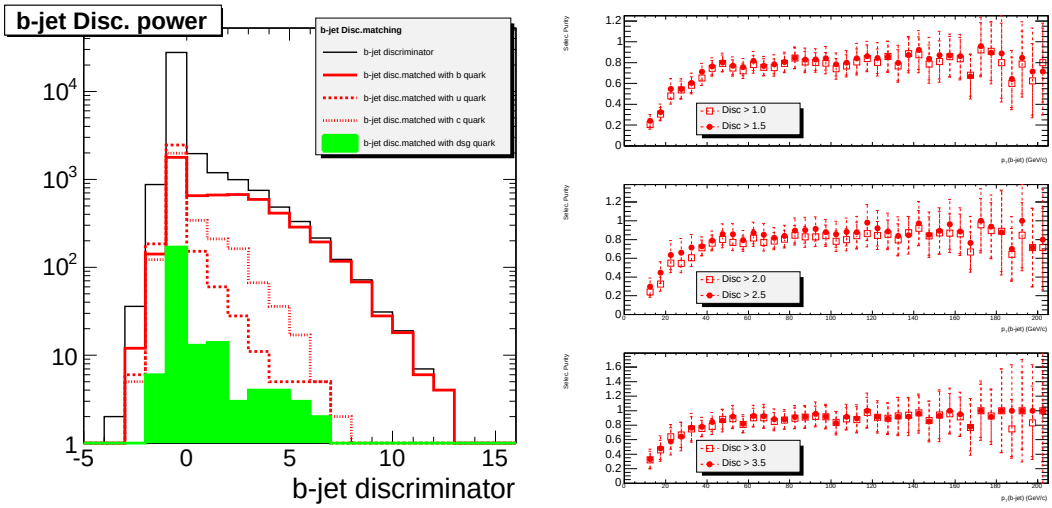

Figure 3.6: Left: The distribution of the discriminator variable output by the "combined b-tagging" algorithm for u/c-tagged jets, for b-tagged jets and for dsg-mistag, as determined from generator-level MC information in the signal with $t \to qZ$. More than 2 tracks are requested, thus most of semileptonic decays end in the left part. Histograms are overlaid, not stacked. Right: The fraction of tagged jets matched with b generated quarks, as a function of p_T(b-jet) and for different discriminator thresholds. Errors are the quadratic sum of statistical uncertainty on each bin.

As a preliminary check, the distribution of $Disc$ from all jets is compared with the one for jets originating from different quark flavour. The match is performed by assigning a flavour to the quark originating from top, staying at a close distance $\Delta R < 0.3$ from the jet direction; the jet is requested to have more than two tracks. The match is repeated for b, c, u quark (that can be produced by the top decay) and from d, s quark and gluon (g) that should be rejected as best as possible. The discriminator is produced for all these cases and the comparison is shown in Fig. 3.6(left).

The bulk of the distribution includes most of the dsg component, together with the events where the quarks decay semileptonically ($b \to ql\nu$, $c \to ql\nu$) or where a u quark hadronizes in a $\pi^{\pm}$ decaying in $\mu\nu$. The invisible neutrino hinders the reconstruction of secondary vertexes, so in these case the b-tagging cannot work (at least a soft-lepton tagging method is explicitly adopted). For higher $Disc$ values the probability to tag a non-b-jet is weaker and weaker, and the match with b-flavour is very good for $Disc > 5$. As expected, discriminator demonstrates to be a powerful variable to tag the b-jet. A lower threshold about $Disc = 2.0 \div 2.5$ should be reasonable, because it allows the containment the c contribution under 10%, the u one under 1% and the dsg under some permil, without major losses in efficiency. The minimum number of tracks in the b-jet candidates has been fixed to 2.

From the same comparison, it is clear that $Disc$ is not optimal to identify a non-b quark. When $Disc < 1$, all the contributions build the distribution and the u/c quark from the anomalous decay cannot be found via an anti-tagging. In addition, u quark are more abundant than c in this region and the sensitivity of the analysis would be biased toward the $t \to uV$

decay. Hence the *Disc* resulting un-useful for the light-jet search.

Though only the signal with $t \to qZ$ is shown here, no different behaviour is expected from discriminator in the sample with $t \to q\gamma$. Considerations above are assumed to be reasonable also there.

Performances of b-tagging

The signal $t\bar{t} \to (qZ)(\bar{b}W)$ is again considered as a case study, to analyze the contamination of light-jets in the b flavour selection based on a discriminator. The 'purity', intended as the fraction of tagged b-jets matched with b generated quarks, is plotted in Fig.3.6(right) in function of jet transverse momentum. Discriminator thresholds defining the b-jet are varied and the statistical error is represented.

For all the *Disc* choices, purities stabilize around 80% when p_T is high. This was expected, as contaminations to the b-jet sample mostly arise from low energy hadrons, that are much more likely to have light flavours. For a given p_T threshold, the purity reasonably increases with the minimum discriminator imposed. The price of increasing the *Disc* cut to values greater than $3.0 \div 3.5$ is a strong loss in efficiency, as too many b-jets with lower values are missed. Once discarded the *Disc* under 2.0, where contaminations from other flavours is too high, the better statistic is found when the the minimum *Disc* is between 2.0 and 3.0 and the p_T greater than $40 \div 50\,\mathrm{GeV}/c$, that allows a 80-85% purity. Even higher purity values can not be obtained, since as Fig. 3.7(right) shows a percent contribution from c-jets is unavoidable.

The non-b-jet component in this selection may arise from non-beauty hadrons originating close to the primary vertex, along with badly reconstructed b-hadrons. The most interesting component is the 'mistagging' with the jets spreading from u and c quarks, as if they are tagged as b-jets the signal has a close resemblance with the $t\bar{t}$ and $Zb\bar{b}$. The misidentification with these flavours has been measured and the result is shown in Fig. 3.7(left) for the contamination with the u quark and Fig. 3.7(right) for the contamination with the c quark. The fraction of jets from u and c quark, wrongly tagged as b, is plotted as a function of p_T and discriminator.

In much of the working space this mistagging can be considered safely low. The small raising of the u component for low momentum may be ascribed to the parton content in the proton and is ruled out by the p_T cut. On the other side, hard light-jets are the result of the anomalous top decay. Among them, some of the more boosted jets containing c quarks (directly from top or from fragmentation) may have a secondary vertex well displaced, challenging the discrimination capability with the b-jets. As expected, this effect is reduced with strengthening the *Disc* cut. The effect is not mitigate by adding more severe cuts (as more charged particle tracks in the jet), and an upper threshold on transverse momentum will deplete too much the signal. On the other hand, this effect is important only at high energy, so the mistagging with u/c integrated between 20 and $200\,\mathrm{GeV}/c$ is below 10% for c

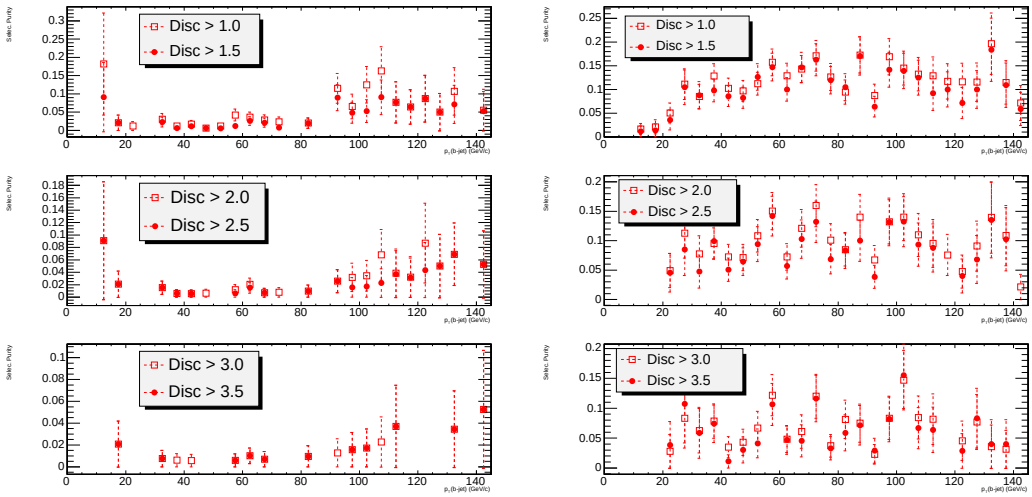

Figure 3.7: The mistagging of a b-jet with a jet coming from a u(left) and c(right) quark, intended as the fraction of tagged jets matching the light quark, for six reference discriminator thresholds. Although the low statistic induces sizable fluctuations, when $2.0 < Disc < 2.5$ the mistaggings are below 15%.

Table 3.11: The contribution of a quark different from b in the devised tagging procedure, as a function of the lower discriminator cut. These estimates are obtained by integrating on the whole p_T spectrum and statistical error is on the last digit.

$Disc$	> 1.0	> 1.5	> 2.0	> 2.5	> 3.0	> 3.5
b-Mistag with c	0.118	0.106	0.096	0.080	0.062	0.055
b-Mistag with u	0.023	0.019	0.015	0.014	0.011	0.009

and below 2% for the u quark for the best discriminator cuts. Small statistic at high energy induces the wide error bars. Table 3.11 summarizes these mistagging estimates for the six different $Disc$ thresholds.

The sensitivity to the low-energy multi-jet component (*i.e.*, the spurious contribution in Fig. 3.6(left) is below 1%.

Once the selection has been proven to have a good purity and a well-controlled 'impurity', the b-tagging efficiency is worth to be measured from simulated event samples. In each p_T bin, this is evaluated as the number of b-tagged jets divided by the total amount of generated b-jets coming from the top decay. Since for low energies the most of tagged jets are not the ones from t, the ratio have to be scaled from a relative purity. Such purity is in turn evaluated as the fraction of tagged jets close to the b quark, sorted within all the jets close to the the generated b and having the same kinematical cuts. Figure 3.8 reports the result of such evaluation, again in function of transverse momentum and for different discriminator cuts.

As expected, with raising the discriminator the criteria for b-jets satisfying the tagging

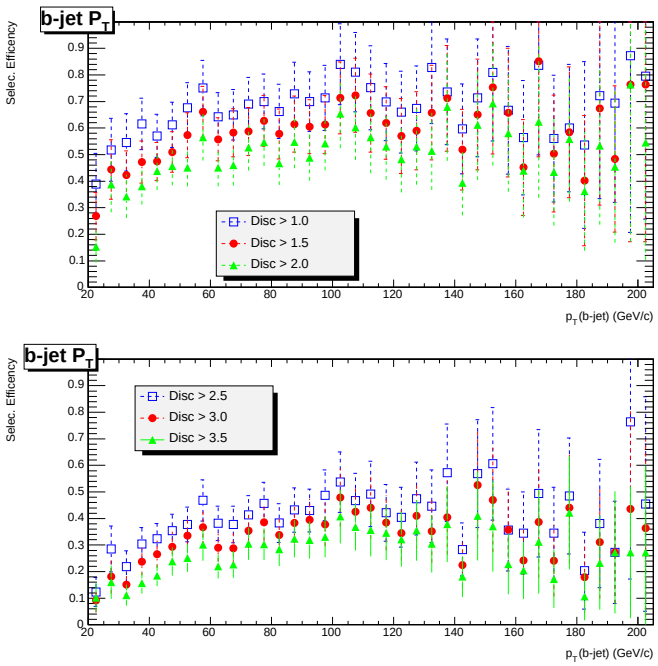

Figure 3.8: The selection efficiency for the b-tagging procedure, as a function of transverse momenta and estimated for different discriminator cuts. When these results are combined with observations on the sample b-purity, the choice for $Disc > 2.0$ turns out to be optimal.

become more severe, thus the percentage of tagged objects decrease. With the values of jet p_T lower cut earlier envisaged, the $Disc$ threshold that prevents the efficiency to fall under 50% is $Disc = 2$. The optimal purity is obtained when $p_T(b\text{-jet}) > 50\,\text{GeV}/c$. The final result of this set of choice is a b-tagging efficiency in the $50 < p_T < 250\,\text{GeV}/c$ range of $(51.5 \pm 2.1)\%$, where the error is statistical only.

As often explained, electrons or muons coming from b quark semileptonic decays have the majors responsibilities in making the background to pass the lepton selection. Therefore, the described selection is also used to find leptons close to the b-jets ($\Delta R < 0.02$), flagging them and then subtracting from the hard leptons sample.

Light-quark jet selection

Distinguish the non-b, hard jets coming from the FCNC top decays is another challenge of the analysis. To cope with it, b-discriminator has proven to be not a powerful variable: therefore, all the calibrated jets reconstructed trough the tracker and the calorimeters are taken into account. Firstly, they are 'cleaned' from effects left by objects already reconstructed in the analysis. If a jet is matched in space ($\Delta R < 0.1$) and energy ($\Delta E_T/E_T < 0.1$) with the track of an electron, the cluster from a photon or a selected b-jet, it has little chance to be a genuine jet from top, so it is discarded. Once these pre-selections have been applied, the objects is requested to not coincide exactly with reconstructed b-jet ($\Delta R < 0.02$, $\Delta E_T/E_T < 0.02$). To

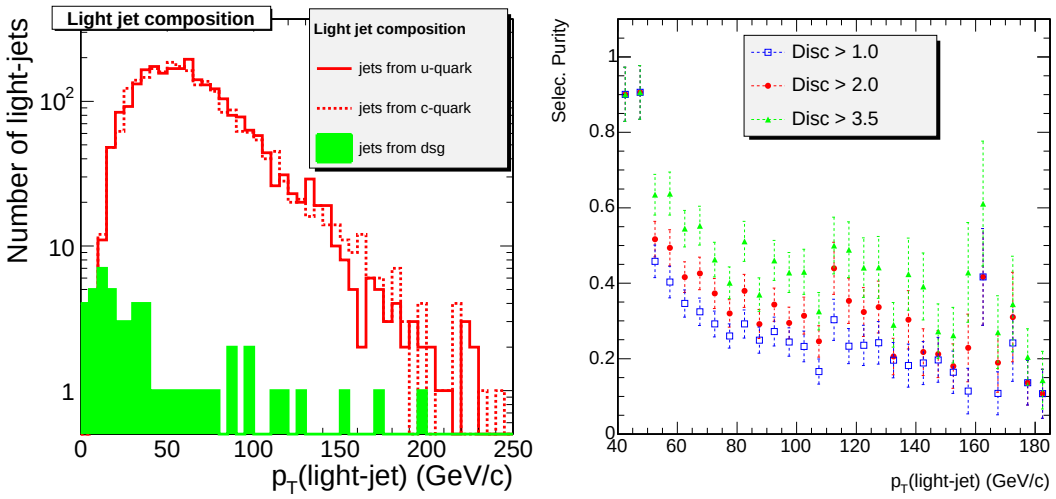

Figure 3.9: Left: A study of the contents of selected light-jets. Two equal parts are matched with the u or c quark from top anomalous decay all along the spectrum, while the fraction of spurious quark is well suppressed above some tens of GeV. Right: Relative mistagging of a light-jet with a b-jet. The high value for these inefficiencies is intrinsic to the jet selection algorithm, and cannot be further minimized.

inquire the origin of these light-jet-candidates, the cone around the reconstructed jet direction is explored and the generated quarks close to it ($\Delta R < 0.3$) are looked at. The transverse momentum distribution of such 'matched light-jet' is reported in Fig. 3.9(left) for the jets coming from a u or c quark from a top and from dsg partons with different origin. This result shows that, already at this level, jets from the minimum bias processes are successfully rejected and no preferences for jets coming from u rather than c are exhibited.

While the contribution from the other jets are sizable below $30 \div 40\,\mathrm{GeV}/c$, where differently flavoured hadrons can be created from the parton sea, in much of the p_T spectrum these hadrons have no energy enough to compete with the jets from top. After about $150\,\mathrm{GeV}/c$, few energetic objects from the signal survive and contaminating jets are no more an issue.

The plots in Figure 3.10 compares the p_T (left) and η (right) of the pre-selected objects with the matched jets and shows that, for low energies and large pseudorapidities, it is not possible to tag these jets in an affordable way. In the first tens of GeV, QCD effects as flavour excitation, gluon splitting, final state radiation completely dominates and the top anomalous decay is overwhelmed. A big amount of these processes occurs at $|\eta| > 2.5$, where jet identification relies on the HF detectors and few of the interesting events are supposed to lie.

Kinematical cuts are the most natural mean to get rid of these contaminations. After retaining only jets with $|\eta| < 2.3$, a selection purity is measured as a function of the calibrated jet p_T, and results are displayed in Fig. 3.11. The purity is intended as the fraction of these jets that are matched with the u, c, dsg.

The plot on the left shows the fractions of selected jets matching with the interesting u

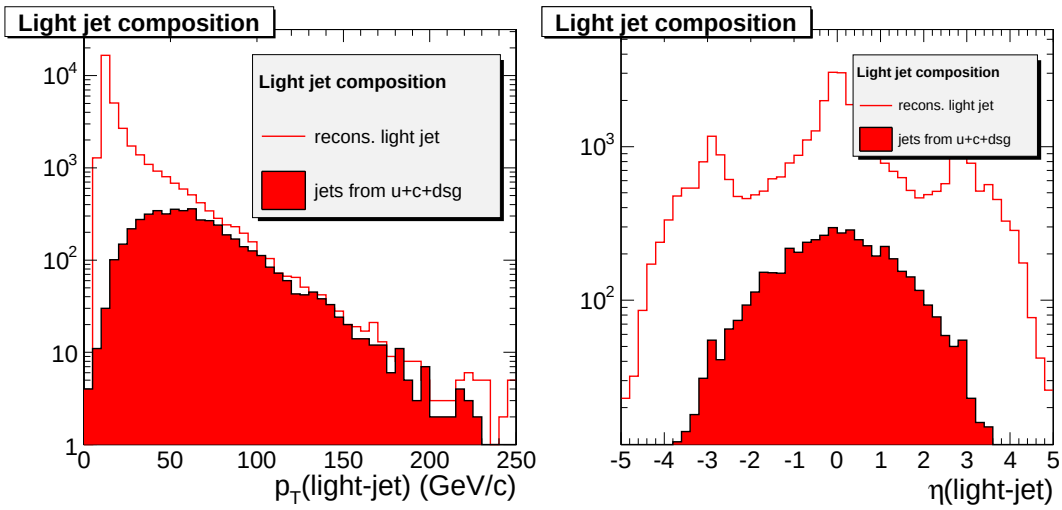

Figure 3.10: A comparison between reconstructed jets (open) and jets matched with u and c quarks from top (filled area), as a function of jet p_T (left) and η (right). Only a good isolation, a pseudorapidity cut and a hard selection can help to reduce the impact of jets from a mass of other sources.

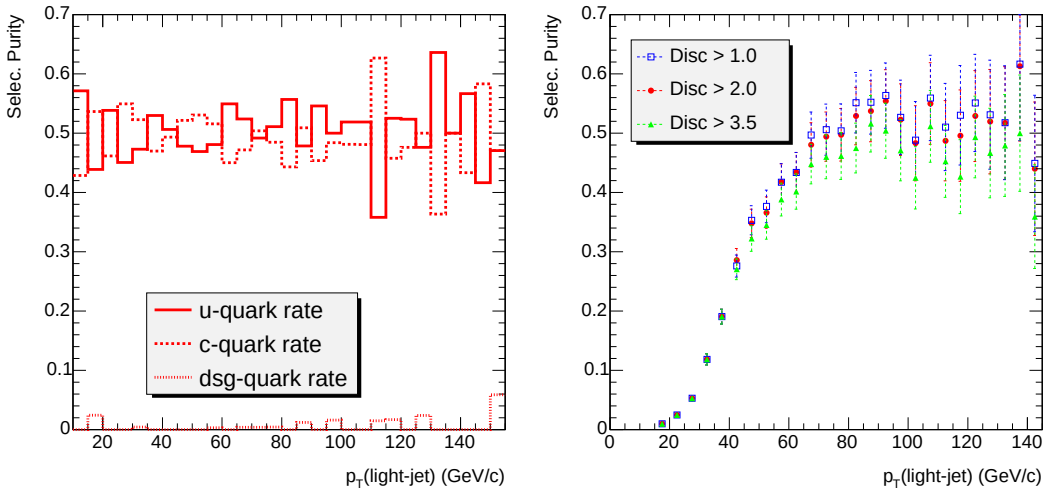

Figure 3.11: Left: The rate for light-jets matched with u, c or dsg quark shows that mis-identification of quark from top is very well controlled and there are no preferences for one flavour of the FCNC decay. Right: the purity of the selected light-jets in function of their p_T. Dependences from the discriminator are a relic of the b-jet subtraction from the light-jets sample.

and c quarks, along with the contamination with objects spreading from the dsg partons. A nice complementarity between the two target flavour is confirmed, since when a jet is not matched with u it is with c quark and vice versa. This is important, because it indicates that the analysis has no different sensitivity between the $t \to cV$ and the $t \to uV$ decay channels within the statistical uncertainty. Even if the occurrence of the u emission is expected to be suppressed with respect to the c one, here it is assumed that the equalization of responses is the best choice for optimization. In fact, maximizing the sensitivity to the only c-jet would not improve the signal efficiency significantly, while only a reduced part of the final-state with an important contamination from u would be accessible.

The contribution of different flavours is contained below few percent after a 40 GeV/c momentum.

In the plot on the right, the selection is showed when different discriminator cuts for the b-jets are adopted, to see whether a preferred value is indicated. Since this variable enters only the jet cleaning from tagged jets, no strong dependence of purity is exhibited. The samples associated to lower $Disc$ cut show a slightly higher purity, because a larger fraction of objects is subtracted. The fact that this purity saturates around $50 \div 55\%$ indicates that other flavours are likely to enter the sample, so this constitutes a potential drawback. Since the jets raising from dsg partons have been completely eliminated in the selection (and no sources of u/c quark different from the top exist in the signal), the spurious contribution may come only from b-jets. To verify the hypothesis, the fraction of selected jet strictly matching a b quark ($\Delta R < 0.02$) has been measured in function of p_T(light-jet) for three different discriminator choices. Figure 3.9(right) confirms the hypothesis, showing that even for a large acceptance in $Disc$, a significant part of b-jets is not subtracted. Only when jets energies are rather high the non-b-contribution is dominating, and there are few chances to tag a further bottom quark where searching for a light one. Consequently, this effect dictates the choice of the p_T lower cut. To contain the light mistagging without loosening the efficiency too much, the 60 GeV/c threshold is envisaged.

In the following sections, performance of these selections in presence of background will be evaluated and further optimized.

3.3 The FCNC $t \to Z^0 q$ decay channel analysis

When searching for a $t \to Zq$ signal, two major issues must be considered: the selection of the three leptons coming from W and Z decays and the discrimination between the light-quark and the b-jets. A strategy for addressing these issues, while ensuring high efficiency and low background contamination is outlined below.

3.3.1 Trigger and Lepton pre-selection in $t \to Zq$ signal

The only events considered in this analysis are the ones in which the $W^{\pm}$ and Z^0 bosons decay leptonically. Therefore, the "double electron or double muon" trigger criteria at the

Table 3.12: Trigger efficiencies (in percent) for the signal and the complete set of considered background sources. Errors are statistical only.

Selection	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Zb\bar{b}$	$ZZ \to 4l$
L1	90.1 ± 0.3	47.97 ± 0.07	99.7 ± 0.3	60.9 ± 0.3
HLT	89.0 ± 1.3	42.95 ± 0.07	92.3 ± 0.3	59.1 ± 0.2
Selection	$Z(\to ll) + j_{85-150}$	$Z(\to ll) + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
L1	37.92 ± 0.15	45.66 ± 0.18	16.81 ± 0.19	26.41 ± 0.11
HLT	36.01 ± 0.15	43.20 ± 0.17	16.18 ± 0.19	24.81 ± 0.16

L1 and HLT step is a natural choice.

In the HLT, isolated leptons are only considered if they have passed the pre-defined trigger thresholds. In the case of the di-muon trigger, both muons must satisfy $p_T > 7\,\mathrm{GeV}/c$, while in the di-electron case, the requirement is $p_T > 17\,\mathrm{GeV}/c$. Acceptance efficiencies are at L1 and in the HLT, for the signal and the full set of background sources is reported in Tab. 3.12.

Electron pre-selection in $t \to Zq$ signal

One of the first step of this analysis consists in applying some of the selection criteria presented in Sec. 3.2.1 to the objects sorted by the trigger.

Figure 3.12(left) compares the transverse momentum distribution of the triggered electrons (thick solid line) with the 'true' information from simulated events. It shows that the electrons coming directly from the decay of a b quark are still significant at this level, especially for the lowest energy, going from some 15% of the selected sample for $p_T > 15\,\mathrm{GeV}/c$ and remaining around 5% even at the highest p_T. Furthermore, this is not the only source of spurious electrons: the gap between the sum of vectorial bosons decay ($e \leftarrow W/Z$) and b quark decay has to be filled with soft electrons from photo-conversions, cascade decays ($b \to c \to e$) or charged pions ($K^{\pm}$, $\pi^{\pm}$) mis-interpreted as electrons.

Furthermore, there is still a large room for electrons from background events, namely W/Z boson, di-boson and multi-jet production. In order to improve the purity of the selection and reduce the background contribution, some among variables related at the electron detection and listed in Sec. 3.2.1 are observed, for the $t\bar{t} \to qZ + \bar{b}W$ signal and the relevant background.

Since at the last cuts of the whole selection algorithm the important processes are only the $Zb\bar{b}$ and the $t\bar{t} \to bW + \bar{b}W$ production, the electron pre-selection will be optimized comparing the signal just with these kinds of events. Optimization is performed looking at the maximum value of the $N_S/\sqrt{N_B}$, where $N_S(N_B)$ is the number of selected signal (background) electron. In the next chapter, this estimator will be demonstrated to be appropriate for the goal. In the following plots, all distributions are normalized to 1, aiming to estimate the discrimination power of the selection irrespectively from cross section of each process.

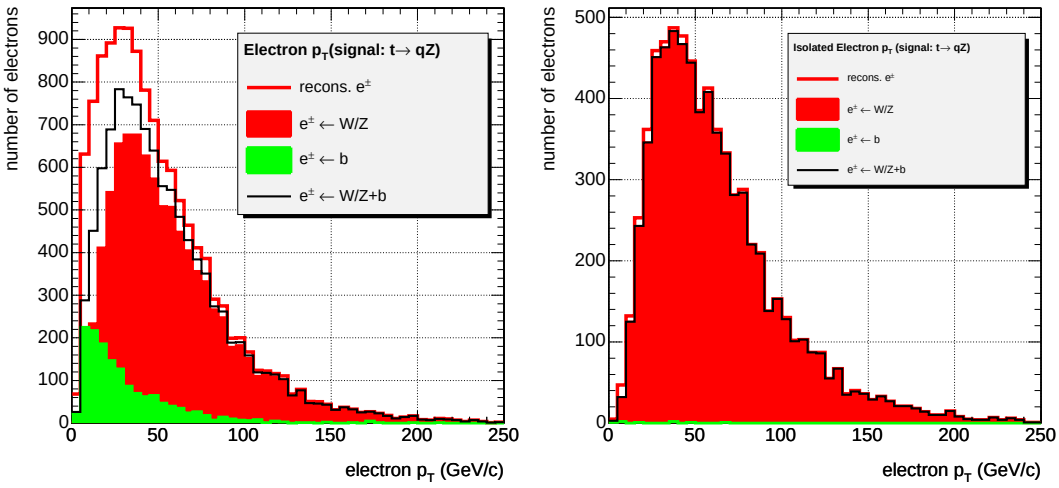

Figure 3.12: The p_T of reconstructed electrons in the $qZ + bW$ signal (thick solid), compared with the distribution of the true electrons. Comparison is shown with true electrons from W and Z (red or dark grey, b-jet decay (green or pale grey) and the sum of two (thin solid). The gap between the sum and the reconstructed electrons has to be filled with the contribution from jets, c-decay and photo-conversions. No isolation are applied in the left plot, while the right plot is for pre-selected electrons.

Isolation is known to be among the most powerful variables for electrons. Figure 3.13(left) shows the distribution of $Isol(e)$, that has been defined in Sec. 3.2.1. It is straightforward to observe how the presence of a couple of b-jets (featuring and average energy lower than those from $t\bar{t}$) reflects on higher values of the isolation variable. Therefore, a cut with $Isol < 0.08 \div 0.1$ (confirmed by the optimization) is instrumental in rejecting the track halo accompanying the electrons from b-jet. On the other hand, electrons from signal and $t\bar{t}$ are expected to share a very similar behaviour, so the separation from this background will follow a different strategy.

Figure 3.13(right) is instead the distribution of the ratio E_{rec}/p_{in} between the energy deposited in ECAL and the corresponding momentum measured in the tracker. As anticipated above, for well-reconstructed electrons this ratio is expected to lie about in the $0.9 \div 1.2$ range. The analysis confirms two narrow distributions for signal and $t\bar{t}$, while the one for $Zb\bar{b}$ is quite larger. The effect is still imputable to hadron activity mixing with the electron signature: for a given track matching the electromagnetic cluster, the larger amount of energy is not deposited in ECAL.

The largest $N_S/\sqrt{N_B}$ is obtained by selecting $E/p > 0.8$, even if a finer optimization will be performed later. On the other hand, upper limits on E/p have demonstrated to not reject the background without affecting too much the signal efficiency. Thus a relaxed $E/p < 100$ can be chosen, just to exclude the largest energy deposits in ECAL with low-p_T tracks, that have to be ascribed to very big bremsstrahlung losses and "showering" electrons.

The E_{HCAL}/E_{ECAL} is a quantity that helps in separate the signal from the $Z + b\bar{b}$ back-

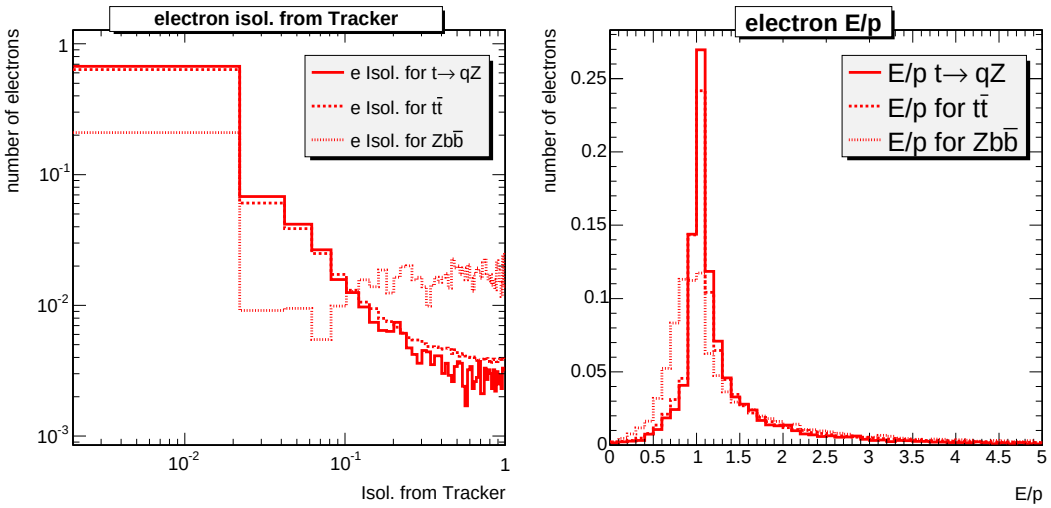

Figure 3.13: Isolation in the tracker (left) and E/p for electrons in signal, $Zb\bar{b}$ (dotted) and $t\bar{t} \to Wb + W\bar{b}$ (dashed) production. A normalization to 1 is applied to all the distributions.

ground. While most of electrons from W and Z decay have no or little energy in the hadron calorimeter (so saturating the first bin of the distribution), those from b quark decay are hidden inside jets, so their identification is hampered by objects releasing a significant fraction of energy in HCAL. From Fig. 3.14(left) one can argue that a cut about $E_{HCAL}/E_{ECAL} < 0.04$ could have an effect even toward the standard $t\bar{t}$, that produce one b-jet more than the signal so it should be characterized by a larger E_{HCAL}/E_{ECAL} for electrons.

Optimization shows that adjusting the E_{HCAL}/E_{ECAL} upper limit in the range $0.02-0.04$ results in a variation of less than 1% in the rejection power. Therefore, this variable will be no further optimized and $E_{HCAL}/E_{ECAL} < 0.03$ will be the default cut.

The last pre-selection variable considered is the matching in η between the candidate tracks and the associate cluster in the electromagnetic calorimeter. It could measure a mis-matching, due *e.g* to photons and neutral hadrons close to the electron, producing energy but without associated tracks. This typically happens for not isolated candidate. Figure 3.14(right) shows that the the bulk of $|\Delta\eta|$ for a hard electron is found in a $\Delta\eta \sim 0.005$ around the track, while the distribution widens for the $Zb\bar{b}$ signal.

The optimal choice indicated from these distribution, that is useful to further reduce the $Zb\bar{b}$ contribution, is $|\Delta\eta(\text{tk-Cl})| < 0.0032$. This further cut is added to the pre-selection parameters.

Performances of the electron pre-selection and momentum cut

In Figure 3.15 the purity of the electron selection is reported, *i.e.* the ratio between the objects matched with the true electrons and the reconstructed one, as a function of both p_T (left) and η (right). As clearly visible, it strongly benefits from the quality cuts. While before pre-selection the percentage of 'true' electron was only $\div 60\%$ after a p_T cut at $20\,\text{GeV}/c$,

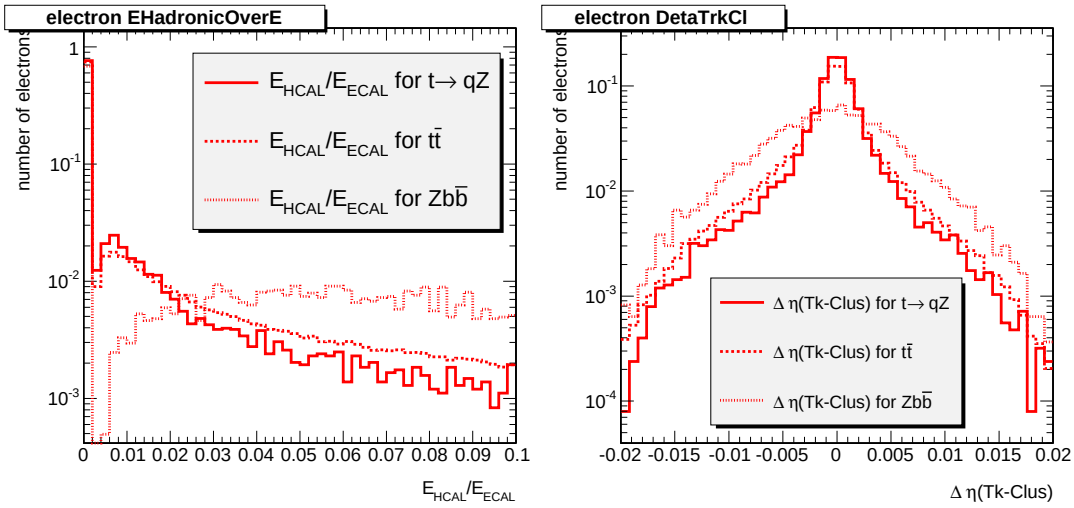

Figure 3.14: The ratio E_{HCAL}/E_{ECAL} between energy in HCAL and ECAL for the candidate electrons (left) and distribution of the distance along η between the cluster and its track (right). Discarding objects with large value of E_{HCAL}/E_{ECAL} (likely to come from hadron decays) and restricting to a narrow window around $\Delta\eta = 0$ allows the improvement of the signal purity.

reaching 80% only after 80 GeV/c, after isolation and other cuts it goes fairly to 95% after the momentum threshold, and stabilizes to a 98% for p_T >40 GeV/c.

The purity is nicely uniform along the whole η range, and an overall improvement about $30 \div 50\%$ is expected. Hereafter, the erros bars are obtained by a quadratic sum of the counting errors in each bin. The high "horns" at the sides for the not-isolated distribution have very large error bars and are due to the lack of statistic in this region. They are recovered after the isolation.

In all, 'contaminating' electrons due to photo-conversions, QCD and heavy-flavour decays are reduced to a minimum. The plot in Fig. 3.12(left) can be repeated after this selection and a figure as Fig. 3.12(right) is obtained. Almost all the reconstructed electrons come from W or Z, thus fully meeting the goal; the contribution from b-decays (indicated as $e \leftarrow b$) reduces to less than 0.001 after 15 GeV/c.

The price to be payed for such a good reconstruction properties is a decrease of the selection efficiency, *i.e.* the fraction of the positively-identified electrons among all the interesting ones. In Fig. 3.16 the efficiency for a single-electron reconstruction (after the double-electron and double-muon trigger stream) is displayed in function of transverse momentum (left) and pseudorapidity (right). The efficiency for electrons from b-decay ("mistag for $e \leftarrow b$"), that is the main objects to be rejected, is also evaluated for different p_T. When p_T exceeds 20 GeV/c, this mis-identification is below 15% and the efficiency rises over 50%. When electron energy approaches to 100 GeV, identification efficiency is as high as $70 \div 75\%$ and mis-identification amounts to some percent. Due to the tiny contribution of electrons from b in the selection, this situation is fully satisfactory.

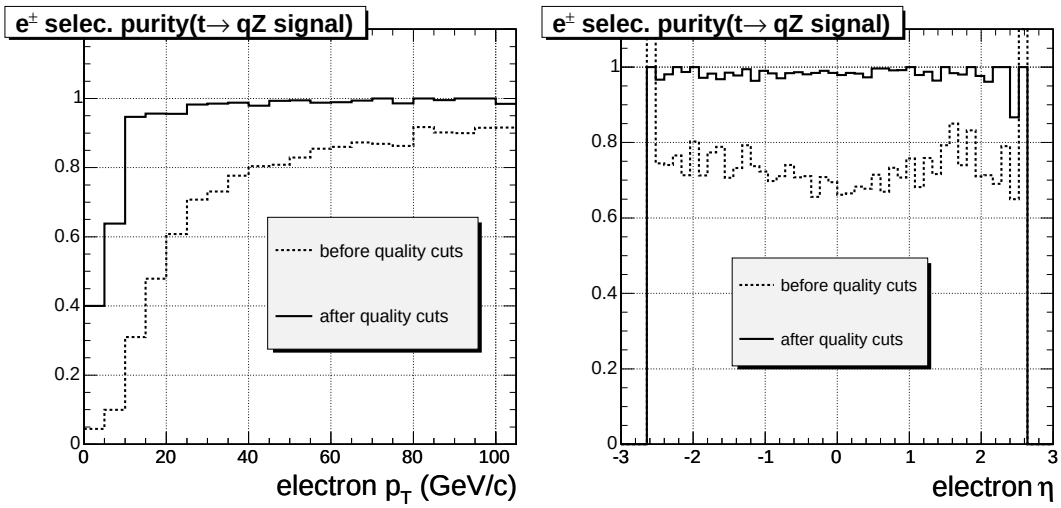

Figure 3.15: The purity of the pre-selected electrons, before and after the quality cuts, as a function of p_T (left) and η (right). The 'horns' at $\eta = \pm 2.5$ are due to a lack of events and larger error bars in this region.

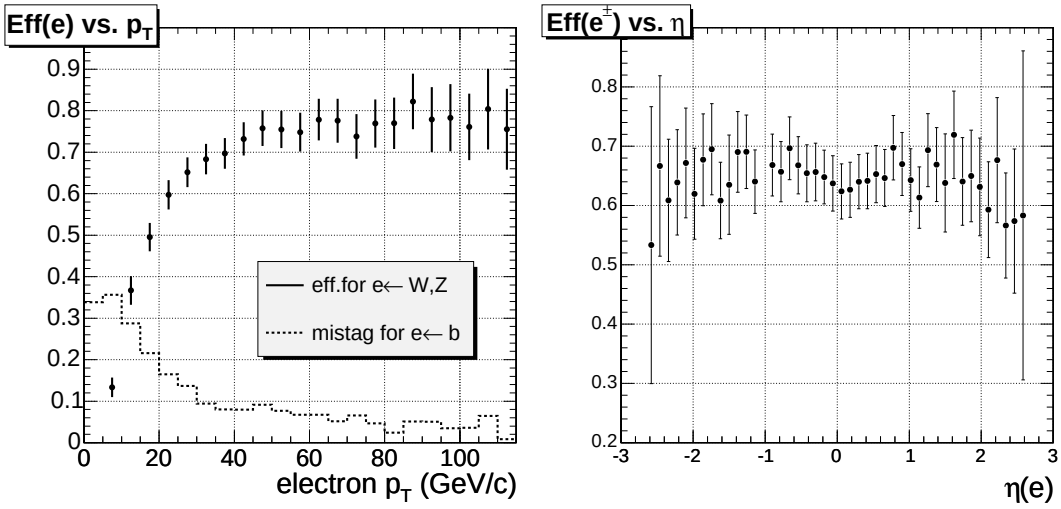

Figure 3.16: The efficiency of the electron selection after quality cuts for the $t\bar{t} \to Zq + W\bar{b}$ signal, as a function of electron p_T (left) and η (right). The error bars are the quadratic sum of the counting error in the histograms used for the distribution, in each bin. The efficiency for 'spurious' electrons coming from b ("mistag for $e \leftarrow b$") is also evaluated.

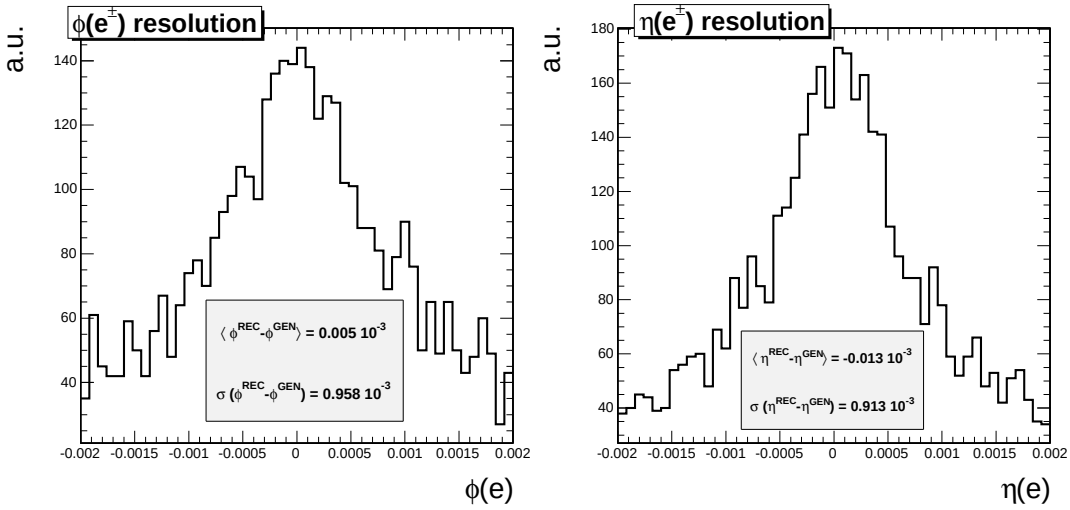

Figure 3.17: The resolution plot for the η (left) and ϕ (right) of the pre-selected electrons, obtained from a difference with the generated (MC) ones. Mean $\langle x \rangle$ and width $\sigma(x)$ are calculated directly from the histograms, without fit overlapped.

The behaviour of efficiency along η, far from measuring the inefficiencies in the different part of the detectors because of the large error bars, has to be considered as an average efficiency integrated over the full p_T range. A separate set of selection cuts in the barrel and in the endcap – that is sometimes adopted in the analyses – does not seem so motivated. As a final comment, it is important to remark that the obtained efficiency is fully consistent with results for other CMS physics analyses (when the effect for different trigger efficiencies is de-convoluted): for instance, it is almost as high as electrons from Z from Higgs (see Ref. [95], Sec. 2.2.2.2), where requirements on electron purity are less tight so quality cuts are looser, and slightly better than electrons from Z inclusive production in the analysis in Ref. [95], Sec. 9.1.2.1.

Another important performance of the off-line reconstruction is the angular resolution in the $\eta - \phi$ plane, as well as the one in energy and transverse momentum. Angular resolutions are obtained from the difference in η and ϕ of the reconstructed and the generated electrons, and are represented in Fig. 3.17.

The mean and r.m.s. of these histograms demonstrate that there are no bias in the angular reconstruction and the resolution is better than 1 mrad for both variables. The tight pre-selection cuts on the angular variable (as those on $|\Delta\eta_{in}|$) largely exceed this value, so they are not smeared by the angular resolution.

The final result from this pre-selection phase can be showed as in Fig. 3.18, where the p_T distribution for the surviving electrons is displayed for the signal and the background. Since the aim of the whole work is to establish an upper limit for the signal, no predictions for the signal cross section are taken into account, and the normalization of the vertical scale is arbitrary. For all the background sources, the number of generated events is rescaled to

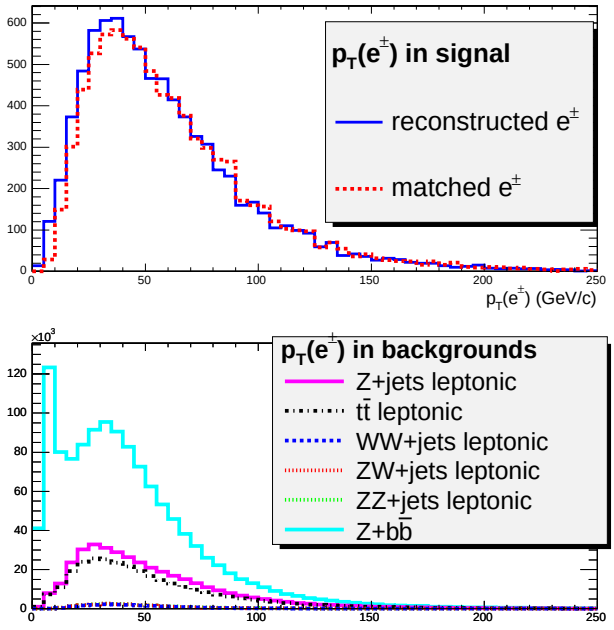

Figure 3.18: The transverse momentum distributions of $e^{\pm}$ candidates after trigger requirements in both the signal sample and the different background samples. The background samples have been normalized to reflect an integrated luminosity of L=10 fb^{-1}, and different classes of events are stacked.

the expected rate after a statistic of 10 fb^{-1} has been integrated, adopting the cross sections presented earlier. The significance of the analysis final result will tell how much of this 'arbitrary units' will be visible on top of background events. This presentation of results will be adopted when displaying all the distributions of reconstructed objects, as electron and muon, vectorial bosons and jets, top quarks.

From plots like Fig. 3.16 it is evident that selection efficiency increase with p_T; on another hand, most of relevant background is not much different in momentum distribution. As a consequence, this threshold needs a careful optimization. In Figure 3.19 the efficiencies of $Wb+Zq$ signal and $Zb\bar{b}$, $Wb+W\bar{b}$ background are compared. Momentum thresholds between 5 and 80 GeV/c are imposed and efficiencies are calculated for each of them. This procedure is repeated for three different upper limits on the one of the most efficient pre-selection variable, *i.e.* $Isol(e)$. Results are displayed for both background processes. Few conclusions can be drawn from these plots:

- while the efficiency drop for $t\bar{t}$ strictly follows the one for signal (being just reduced by a factor $2.3 \div 2.5$), for the $Zb\bar{b}$ there is a region where efficiency for signal can be pushed to a maximum and that for background to a minimum. This corresponds to select electrons with $p_T > 15 \div 20$ GeV/c;
- the strong similarity of electrons in signal and $t\bar{t}$ is also reflected in the dependence from $Isol(e)$, where no definite cuts seem preferred. On the other hand, in $Zb\bar{b}$ the cut

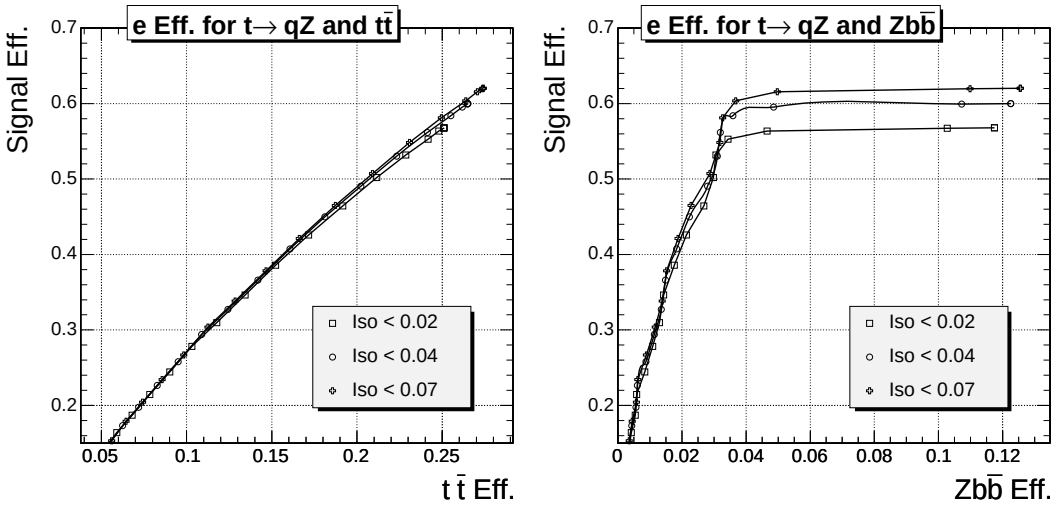

Figure 3.19: The electron reconstruction efficiency for signal vs. the efficiency of the two most relevant backgrounds, obtained for different p_T thresholds. Three possible values of isolation cuts are considered.

$Isol < 0.07$ shows the best efficiency performance.

The conclusion is that the optimized selection cuts are $p_T > 20\,\mathrm{GeV}/c$ (the most conservative for background rejection in the preferred region) and $Isol(e) < 0.07$. A confirmation comes from plots as Fig. 3.20 (left), where the signal efficiency is compared with the 'background rejection', *i.e.* the ratio between signal and cumulative background efficiency for both processes.

A similar procedure is repeated for the E/p variable, in order to finely tuning the lower limit. As Fig. 3.20 (right) shows, there is no significant differences in performance betweens 0.6 and 1.0 threshold, thus the default cut $E/p > 0.8$ is well motivated. As expected, a cut greater than 1 strongly demolishes the signal efficiency. From both plots, it is clear that the best working point is around the fourth open cross marker from the left, that results in a p_T lower threshold of $20\,\mathrm{GeV}/c$. It corresponds to a signal efficiency of 0.520 ± 0.006 and a rejection of 1.807 ± 0.004, that includes an efficiency of $(24.12 \pm 0.10)\%$ for $t\bar{t} \to Wb + W\bar{b}$ and $(3.19 \pm 0.02)\%$ for the $Zb\bar{b}$ production.

The Table 3.13 summarizes all the pre-selection cuts for the electron in the $t\bar{t} \to Zq + W\bar{b}$ signal.

Muon pre-selection in $t \to Zq$ signal

With respect to the the delicate study for the electrons reconstruction, the muon pre-selection is much less demanding. The peculiar CMS design is already enough to ensure that the hard particles crossing the muon stations are almost all muons.

Figure 3.21(left) compares the transverse momentum distribution of the triggered muons

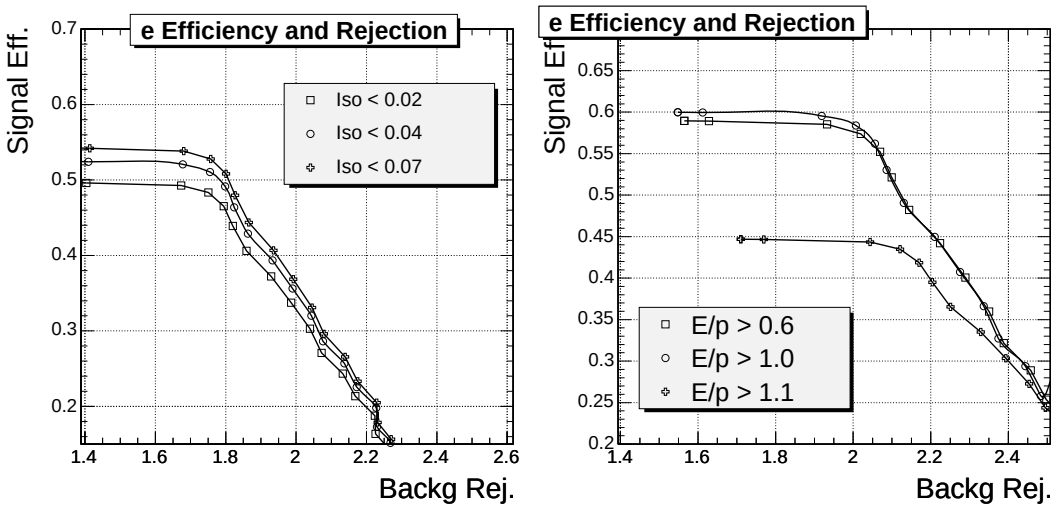

Figure 3.20: The electron efficiency for the signal vs. the background rejection, for different values of $Isol(e)$ (left) and E/p (right) cuts. It indicates the best workpoint, that is around the fourth open cross marker from the left. It corresponds to a p_T lower threshold of $20\,\mathrm{GeV}/c$.

Table 3.13: A summary of electron pre-selection cuts adopted in the analysis. Each choice comes from the search for an optimization and is demonstrated to produce high purity and good efficiency.

L1	Di-electron ("Relaxed") Trigger, 12(19) GeV/c
HLT	Di-electron Trigger, 17 GeV/c
Isolation	$\Delta R < 0.1$, $Isol < 0.07$
Track/Cluster energy	$E/p > 0.8$, $E/p < 100$
Track/Cluster	$\lvert\Delta\eta_{in}\rvert < 0.0032$
HCAL/ECAL	$E_{HCAL}/E_{ECAL} < 0.03$
Transv. momentum	$p_T > 20\,\mathrm{GeV}/c$

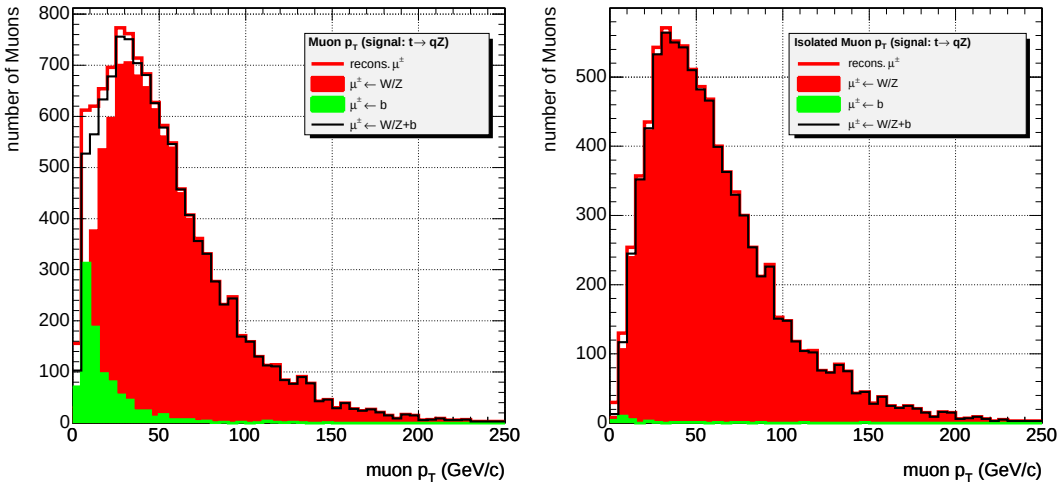

Figure 3.21: The p_T of candidate muons in the $qZ+bW$ signal, as they come from the double trigger, compared with distribution of the true muons. Comparison is shown with true muons from W and Z (red or dark grey), b-jet decay (green or pale grey) and the sum of two (thin solid). The gap between the sum and the reconstructed objects has to be filled with the contribution from jets and c-decay and photo-conversions. No isolation criteria are applied in the left plot, while pre-selected muons distribution is in the right plot.

(thick solid line) with the 'true' information from simulated event samples. The fraction of muons coming directly from the decay of a b quark is significant only at very small energies: when $p_T > 15\,\mathrm{GeV}/c$ it amounts to 4.1% of the total of reconstructed muons and steadily decreases with the momentum.

Even the other contributions (mostly cascade decays $b \to c \to \mu$, that are not supposed to reach the outer chamber but can be reconstructed in the tracker) accumulate close to a threshold of the muon p_T. Nevertheless, some optimizations are possible to improve the purity of the selection.

Best choices for cuts are again found by maximizing the $N_S/\sqrt{N_B}$ ratio performed on distributions normalized to the same number of events. Since the topology of muons closely resembles the electrons one (if both are coming from W or Z), the only background sources taken into account are still the $Zb\bar{b}$ and the $t\bar{t} \to bW + \bar{b}W$ production.

The most useful handle in muon pre-selection turns out to be the tracker isolation ($Isol(\mu)$), that was defined in Sec. 3.2.1 in a cone of radius $\Delta R = 0.3$. Figure 3.22 (left) shows the distribution of $Isol(\mu)$ for the signal and background. Here the muons from b-jets are less prone to be distinguished from the 'genuine' ones, so a threshold for the isolation variable is less straightforward. Numerical optimization shows that, though a cut $Isol < 0.13$ allows to obtain the same $N_S/\sqrt{N_B}$ as $Isol(e) < 0.08$ for the electron case, the maximum can be reached again with a cut at 0.07 as in electron case.

This tight cut provides a significant rejection of the tracks halo lining the muon tracks when muons are embedded inside jets. In order to identify the same contaminating electrons,

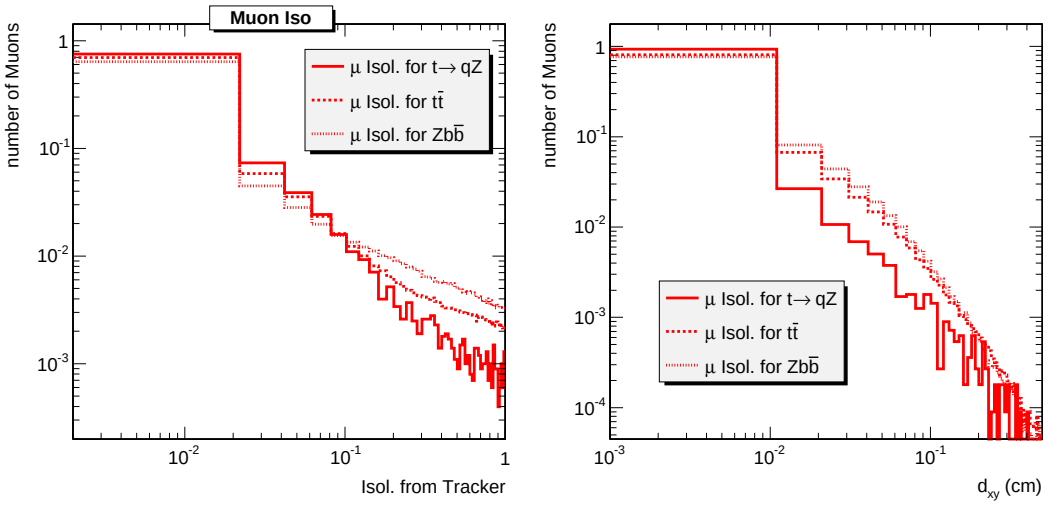

Figure 3.22: Isolation in the tracker (left) and transverse impact parameter (right) for muon from signal, $Zb\bar{b}$ (dotted) and $t\bar{t} \to Wb + W\bar{b}$ (dashed) production. A normalization to the same number of events is applied.

another interesting variable could be the transverse impact parameter d_{xy}, *i.e.* the distance in the transverse plane from the primary vertex and the extrapolated muon track. As Fig. 3.22 (right) shows, both backgrounds exhibit a tail that is longer then the signal (about $d_{xy} > 0.01\,\text{cm}$), that is related to muons coming from a secondary vertex. On the other hand, the bulk of these distributions lies in a quite narrow region, that is loosely sensitive to upper cuts on d_{xy}: therefore, this variable will not be considered for any cut.

The resulting effects on purity and efficiencies of these selections are estimated below.

Performance of the muon pre-selection and momentum cut

In the same way as the electron, the "purity" for the selected muons is defined as the ratio of the 'matched' candidate with the reconstructed objects. In Fig. 3.23 the comparison of this quantity before and after the isolation is shown, as a function of both p_T (left) and η (right). Even before this quality cut, muons from W/Z result to be about 80% of the total if $p_T > 20\,\text{GeV}$; when isolation is required, semileptonic decays contribution disappears and purity exceeds 95% already after $15\,\text{GeV}/c$, reaching a nice 98% thereafter.

The uniformity of the muon purity along η is fairly good and marks a reference value $\sim 98\%$ integrated for p_T from 0 to $250\,\text{GeV}/c$. The minor decrease for $|\eta| > 2$ indicates that muons from b quark are likely to come from region at high pseudorapidity, and it is not an issue here.

Following the analysis for electrons, the transverse momentum distribution in Fig. 3.21(left) can be repeated after the application of the chosen isolation cut. The plot in Fig. 3.21(right) demonstrates how spurious contribution are successfully reduced, even at modest p_T. Contamination is below 1% already at $p_T = 15\,\text{GeV}/c$, falling below some 0.001 for harder

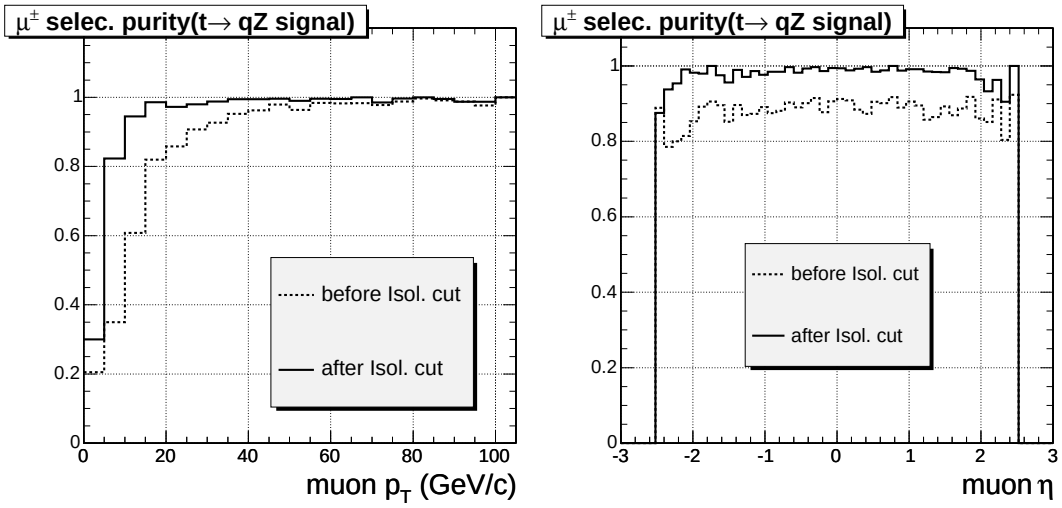

Figure 3.23: The purity of the pre-selected muons, before and after the isolation cut, as a function of p_T (left) and η (right). The 'horns' at $\eta = \pm 2.5$ are due to a lack of events and big error bars in this region.

energies.

The effect on the signal efficiency caused by the quality cuts is evaluated, and resulting plot displayed in Fig. 3.24. The same electron efficiency is reached when the muon transverse momentum is about $30\,\mathrm{GeV}/c$, then it smoothly increases up to $\sim 80\%$. If one remembers that the double electron trigger stream delivers candidate muons with an efficiency not greater than 90% (see for instance Ref. [95], Tab. 10.6 and 10.7) this result – along with the very high purity – seem very satisfactory. Analyses for similar cases (for instance Ref. [95], Sec. 9.1.3 for the Z production) declares quite similar results.

At the lowest momenta after the threshold, misidentification probability with muons from beauty decay is still around 10%, but as shown before the fraction of these muons in the selected sample is totally negligible. A separate set of cuts for the endcap region seem not strongly motivated, though it would need a richer statistic to improve the error bars.

The redundancy of the muon off-line selection system, together with the high confidence that a reconstructed muon is a genuine muon, allows the angular muon resolution to be even better than the electron one. Figures 3.25 reports the distribution of the difference of the η/ϕ reconstructed muon track position with the generated one.

The mean and r.m.s. of these histograms demonstrate that there are no bias in the angular reconstruction and the resolution is better than 1 mrad for both variables. This accurate reconstruction in space for both electrons and muons justifies those selection algorithms based on the geometry of the event, because the angular resolution for leptons is far to be an issue.

The p_T distributions for muon candidates after the application of both trigger and offline quality cuts, in both the signal events and in the different background samples, are represented

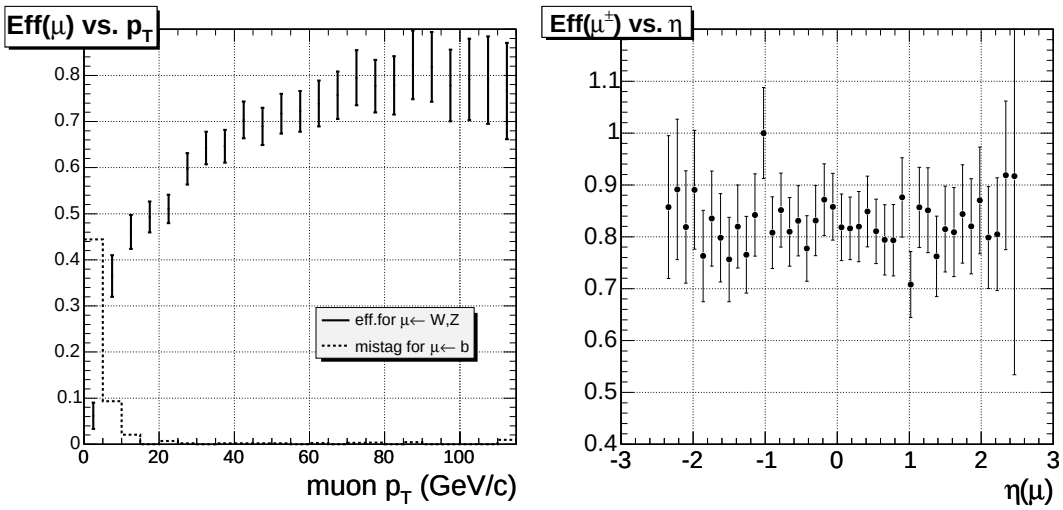

Figure 3.24: The efficiency of the muon selection after isolation cut for the $t\bar{t} \to Zq + W\bar{b}$ signal, as a function of muon p_T (left) and η (right). The error bars are the quadratic sum of the counting error in the histograms employed for the distribution, in each bin. The efficiency for 'spurious' muons coming from b ("mistag for $\mu \leftarrow b$") is also evaluated.

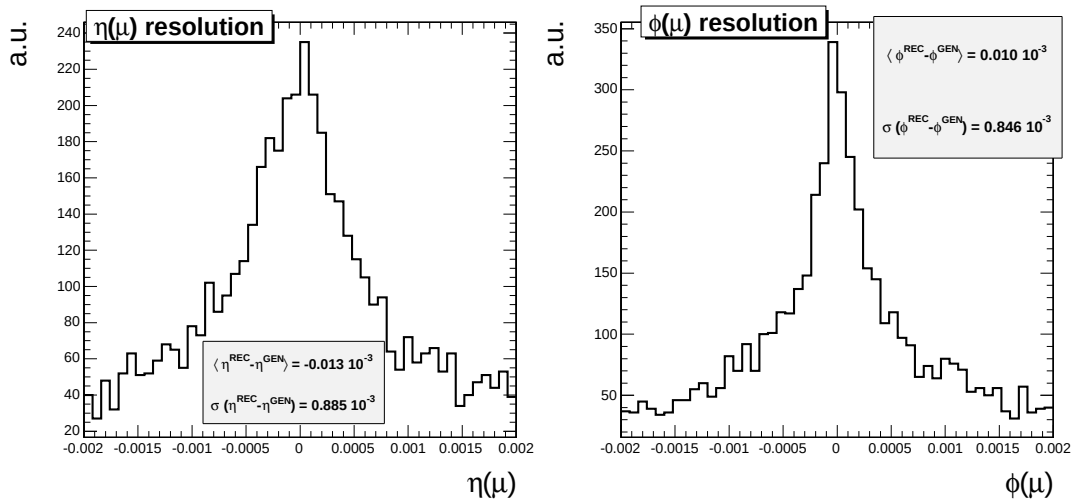

Figure 3.25: The resolution plot for the η (left) and ϕ (right) of the pre-selected muons, obtained from a difference with the generated (MC) ones. Mean $\langle x \rangle$ and width $\sigma(x)$ are calculated directly from the histograms.

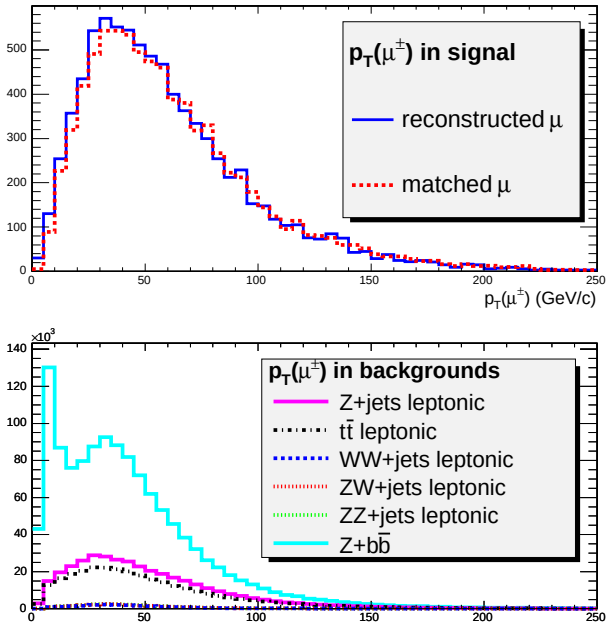

Figure 3.26: The transverse momentum distributions of $\mu^\pm$ candidates after trigger requirements in both the signal sample and the different background samples. The corresponding distributions after the quality cuts have been applied are also shown. The background samples have been normalized to reflect an integrated luminosity of L=10fb^{-1}.

in Fig. 3.26. The transverse momentum threshold could be optimized by issuing a similar procedure to the electron case. Strictly resembling what was done there, signal efficiency is compared with efficiencies for $Zb\bar{b}$ and $t\bar{t} \to Wb + W\bar{b}$, choosing three reference values for the tracker isolation cut and shifting the p_T threshold. As expected, the muon efficiency plot for $Zb\bar{b}$ is strongly similar to that for electron.

In Figure 3.27 the momentum thresholds is varied between 5 and 80 GeV/c and efficiencies are calculated for the same three upper limits on the one of $Isol(\mu)$. Efficiency in the two background samples is found to steeply increase with the p_T cut, in a way that is not different from the signal behaviour. The 'plateau' in the signal efficiency against $Zb\bar{b}$ is not so relevant here. Best efficiency performance is again obtained when the isolation is better than 7%.

From both plots, the best working point is chosen around the third open cross marker from the left, that results in a p_T lower threshold of 15 GeV/c. It corresponds to a signal efficiency of 0.611 ± 0.006 and a rejection of 0.61 ± 0.02, that includes an efficiency of $(30.0 \pm 0.1)\%$ for $t\bar{t} \to Wb + W\bar{b}$ and $(84.3 \pm 0.1)\%$ for the $Zb\bar{b}$ production.

The Table 3.14 summarizes all the pre-selection cuts for the muon in the $t\bar{t} \to Zq + W\bar{b}$ signal.

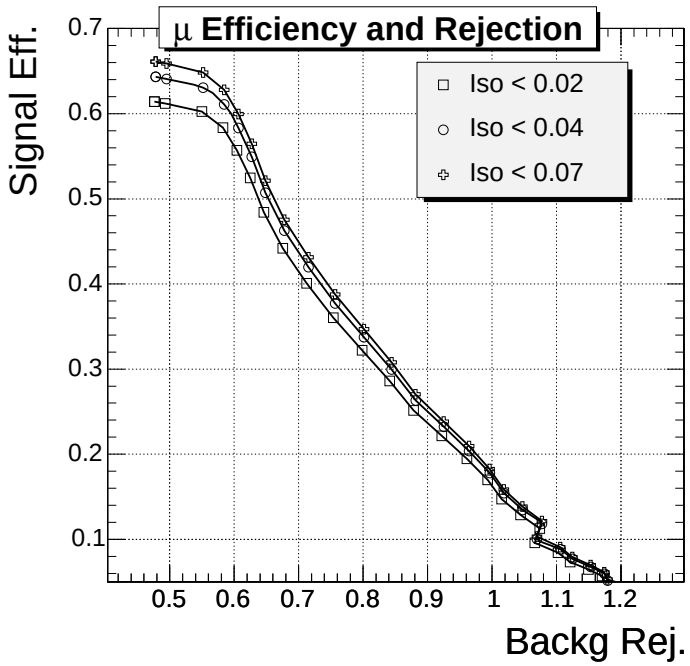

Figure 3.27: The muon efficiency for the signal vs. the $t\bar{t}$ and $Zb\bar{b}$ rejection, for different values of $Isol(\mu)$. The best point is considered to be the third from the left (open cross) and corresponds to a p_T lower threshold of 15 GeV/c and a 0.07 isolation.

Table 3.14: A summary of muon pre-selection cuts adopted in the analysis. Each choice comes from the search for an optimization and is demonstrated to produce high purity and good efficiency.

L1	Di-muon Trigger, 3 GeV/c
HLT	Di-muon Trigger, 7 GeV/c
Isolation	$\Delta R < 0.3$, $Isol < 0.07$
Transv. momentum	$p_T > 15$ GeV/c

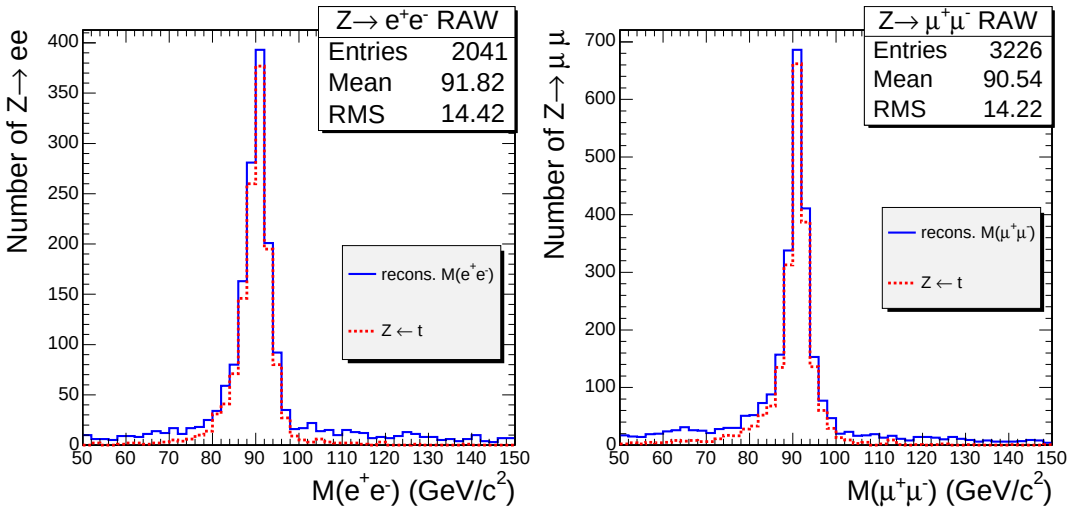

Figure 3.28: The invariant transverse mass of the Z boson, obtained from electrons (left) and muons (right) pairs with kinematical cuts, sorted among the closest to the Z mass. The higher muon efficiency results in quite richer background.

3.3.2 Constraints on the Z^0 mass

In the case of the $t \to qZ^0$ channel, the Z^0 is reconstructed by combining two same-type and opposite-sign leptons. The pair of leptons (muons or electrons) in each event whose invariant mass lies closest to that of the Z^0 is assumed to be correct one. The mass constraint will be applied in most of the invariant mass reconstructions (namely Z, W and FCNC top), since it demonstrates to improve the signal efficiency and better contain the background. Though this technique tends to bias the background, enriching it around the mass constrain and accumulating combinatory effects in the signal region, comparisons with generated 'true' particles from Monte Carlo will comfort about the effectiveness of the reconstruction. In the final state, the combinatorial background will be subtracted by a proper fit procedure.

Among the lepton selection, electrons and muons have to be removed if they are embedded into a b-jet or a light-jet ($\Delta R < 0.5$ from the jet axis), because in that case they probably come from a semileptonic b or c quark decay. From all the same-flavour different-sign leptons pairs, the one with the mass closest to $M(Z^0) = 91.188\,\mathrm{GeV}/c^2$ [24] is sorted. These simple selection criteria may be analyzed separately in the electron and muon cases, by regarding the purity of the reconstruction and the background rejection.

The invariant mass distributions for the $Z^0 \to e^+e^-$ and $Z^0 \to \mu^+\mu^-$ channels, presented in Fig. 3.28, shows the expected asymmetry on the low side of the peak, which is caused by photon radiation. Here the reconstructed boson is compared with a 'matched' object, *i.e.* a Z in which the two leptons are closer than $\Delta R = 0.3$ to the reconstructed one. The typical lineshape is correctly reproduced, with a central value (centroid of the histogram between 50 and 150 GeV/c^2) shifted only some 7% from the pole mass.

Some further cuts could be devised to reduce the lepton pairs from background falling

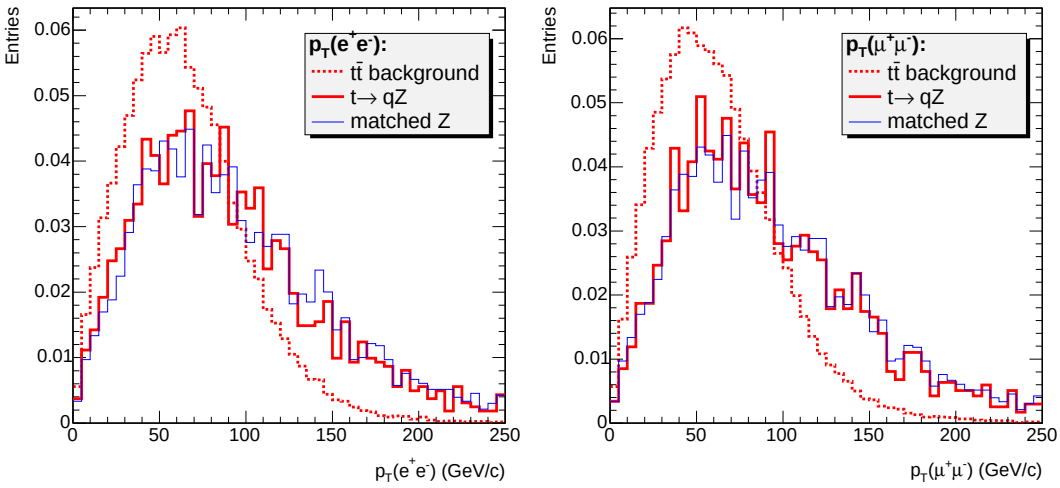

Figure 3.29: The transverse momentum of the reconstructed Z can be evaluated for signal and its background (where lepton pair can combine in a narrow mass window), to check if some further cuts are effective. Here no distinctive differences are exhibited from the two processes, and additional selections are not introduced.

in this mass window. As previously stated, the main concern has to be deserved to the $t\bar{t}$ background, in which leptons from the di-leptonic channel are very prone to imitate the ones from the Z decay. A slower boost is expected from the pairs, so selecting a l^+l^- pair with a some high momentum may reduce this contribution. Figure 3.29 exhibits the distribution of the $p_T(e^+e^-)$ and $p_T(\mu^+\mu^-)$ objects. While the background momentum is significantly slower than the signal one, they are not enough disentangled to suggest a $p_T(Z)$ cut. A numerical evaluation of the $S/\sqrt{B}$, with a running momentum threshold on the whole range, shows that there are no advantages in rejecting Z^0 with low momenta. The transverse momentum for the 'matched' objects is also indicated. Its distribution closely resembles the reconstructed one, thus confirming the quality of the reconstruction.

The requirement of only one reconstructed Z boson is helpful in increase the purity of selection, because it removes the lepton pairs (mostly muons) falling in the Z mass but coming from other sources. The result is reported in Fig. 3.30 and can be compared with the Fig. 3.28 distribution: the tails are cleaned by a 19% and a 25% in the electrons and muon component respectively. The peak resolution results to be increased by a factor $\sim 25\%$ in e^+e^- and $\sim 35\%$ in $\mu^+\mu^-$ and the mass is shifted by only 0.3% for electrons and 1.2% for muons. It is worth to emphasize that a precise pole mass measure is not the main purpose of this search, thus a detailed fit on the distribution is not faced here.

The comparison of the reconstructed Z in the signal (and the matched object overlaid) with all the background is presented in Fig. 3.31(left). The dominant source of background for this distribution comes from the Z^0 plus jets (light or b-flavoured) sample, particularly given their large cross sections. As expected, the introduction of the Z mass parameter in the selection leads to a mighty biased background distribution, with a maximum in the region

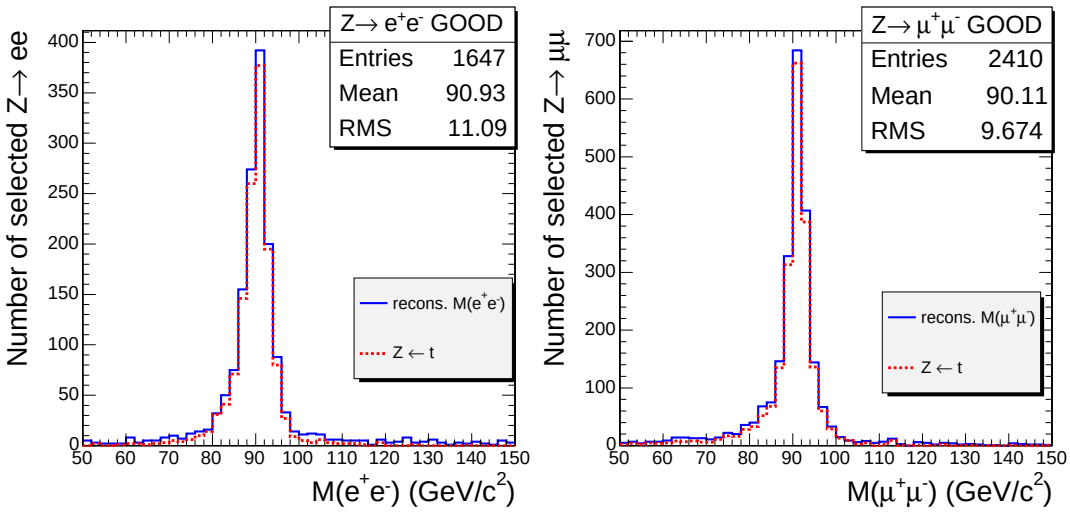

Figure 3.30: The invariant mass plot of the reconstructed Z, under the veto condition for a further Z. The peak resolution improves by about 30%, combinatoric effects reduce and the matching with generated leptons is significantly better.

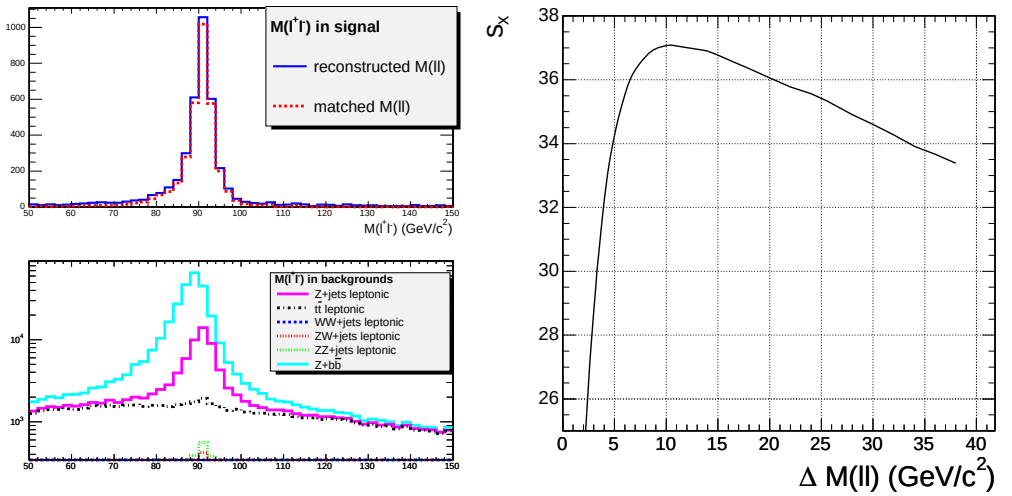

Figure 3.31: Left: The e^+e^- and $\mu^+\mu^-$ invariant mass distributions in the signal sample and in the different backgrounds, where all contributions are showed stacked. The background distributions are normalized assuming an integrated luminosity of L=10 fb^{-1}. Right: The significance of the selection in the invariant mass plot as a function of the window width. The parameter s_x is proportional to the $N_S/\sqrt{N_B}$ and reaches its maximum when $\Delta M\,(ll) = 10\,\mathrm{GeV}/c^2$.

around the Z peak. The spread in the mass peak for the $Zb\bar{b}$ production indicates that lepton pre-selection is not yet insensitive to b quark semileptonic decays: when one of them is mis-identified as a lepton produced from Z, a contaminating effect adds and the mass widens. Not all the considered background sources are represented in this figure and the Z+jets stems for the $Z + j_{85-150}$ component.

One of the most powerful background rejection tool – at least for the non-peaking background – has demonstrated to choose a narrow window around the peak on the invariant mass distribution. This is very efficient in selecting only pairs with the best matching, thus rejecting leptons which do not come from a Z decay. In order to choose the best $\Delta M(Z)$ window, the mass distribution for the signal and relevant background ($t\bar{t}$ and $Zb\bar{b}$) is used to scale the number of selected events. The significance s_x of the analysis, that is related with the ratio $N_S/\sqrt{N_B}$, is exploited here as a simple estimate of the impact of this selection. To not introduce a further bias, the peak mass is left free and the window is opened around the centroid of the reconstructed distribution. This allows to absorb some systematic effects on the invariant mass peak, as the muon and electron energy scale and linearity. From Fig. 3.31(right) it is clear that the best point is around $\Delta M(Z) = 10\,\mathrm{GeV}/c^2$. It is important to observe that this indication of significance is nicely stable with respect to Z mass resolution. In a highly pessimistic case of a 5% electron resolution (and consequently a $\sim \sqrt{2} \cdot 5\% \sim 7.5\%$ resolution in $\Delta M(Z)$), the significance would be lowered less than 0.7%. A corresponding worse in the FCNC branching ratio sensitivity is expected.

A more exhaustive comparison of selection results are offered by Tab. 3.15, where cascade efficiencies after the reconstruction algorithm are quoted for all samples in signal and background.

3.3.3 Missing transverse energy and reconstruction of $W^{\pm}$

The search for a W boson in simulated events is constrained by requiring that the event already contains a good Z^0 candidate. This is very efficient in rejecting processes which have less than three hard leptons, as the WW, Z+jets and $t\bar{t}$ background samples.

Before combining the lepton and the transverse missing energy to form a transverse invariant mass bound, a closer look to the latter is necessary.

Constraints on the Missing transverse energy

As outlined above, the amount of missing transverse energy E_T^{miss} is determined by the Iterative Cone method supplemented by jet corrections. This variable is compared in Fig. 3.32 (left) for the signal and the relevant background.

A similar behaviour is shared among signal and processes producing a W that decays leptonically, while the $Z+X$ production behaves quite differently, making possible to devise a lower E_T^{miss} cut. The missing energy cut is optimized with respect of the $Zb\bar{b}$ ad $t\bar{t}$ background.

Table 3.15: The efficiencies for the reconstruction of a Z boson from a pair of electron or muons, presented for an exhaustive set of background sources. Leptons retain the quality cuts discussed earlier and they are selected in an invariant mass window of $10\,\mathrm{GeV}/c^2$. All values are in percent.

Selection	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Z(\to ll)b\bar{b}$	$ZZ \to 4l$
one good $Z^0 \to ee$, after $\Delta M(ll)$	19.3 ± 0.5	1.460 ± 0.013	32.70 ± 0.13	14.11 ± 0.17
one good $Z^0 \to \mu\mu$, after $\Delta M(ll)$	23.5 ± 0.5	1.500 ± 0.013	34.93 ± 0.13	13.94 ± 0.12
one good $Z^0 \to ll$, after $\Delta M(ll)$	42.8 ± 0.7	1.500 ± 0.013	34.51 ± 0.19	27.08 ± 0.17
Selection	$Z(\to ll) + j_{85-150}$	$Z(\to ll) + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one good $Z^0 \to ee$, after $\Delta M(ll)$	8.34 ± 0.06	10.15 ± 0.08	3.12 ± 0.08	0.357 ± 0.020
one good $Z^0 \to \mu\mu$, after $\Delta M(ll)$	7.41 ± 0.06	9.11 ± 0.08	3.69 ± 0.09	0.440 ± 0.022
one good $Z^0 \to ll$, after $\Delta M(ll)$	15.25 ± 0.09	18.35 ± 0.11	6.81 ± 0.12	0.797 ± 0.29

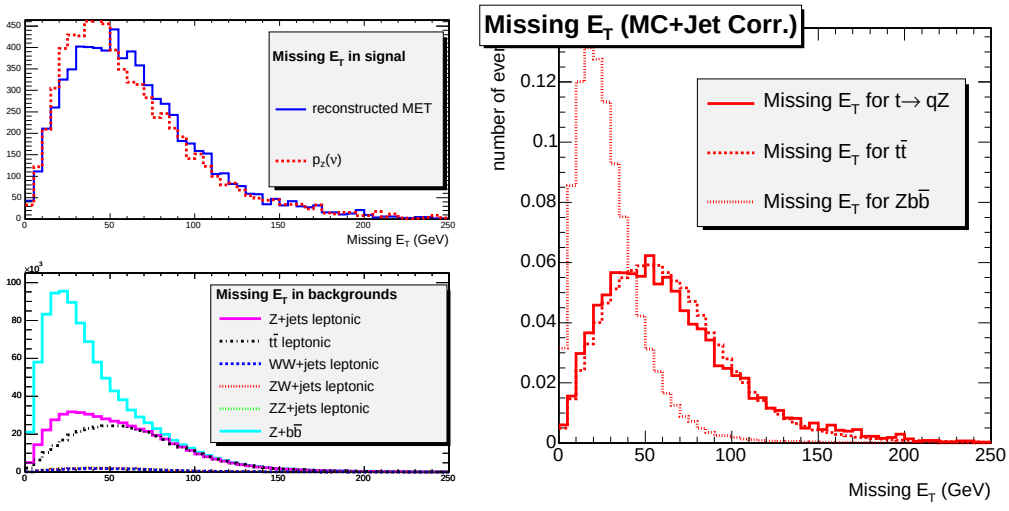

Figure 3.32: Left: The comparison of the E_T^{miss} for the signal and relevant background. In the lower plot, rates are scaled at the $10\,\mathrm{fb}^{-1}$ integrated luminosity. Right: the same distribution from signal and the two most important background sources, when normalized to the same number of events. Since in the $Zb\bar{b}$ channel missing energy comes from b quark decays and τ decays from Z, it is significantly softer and the relevance of $Zb\bar{b}$ can be reduced by cutting $E_T^{miss} > 30 \div 35\,\mathrm{GeV}$.

In Fig. 3.32 (right) they are compared to the signal as they had the same number of events, and the threshold that reflects in the best significance for the final result is searched for. It can be observed that, while $t\bar{t}$ is not distinguishable from the signal at this level (demonstrating that the presence of the additional $W \to l\nu$ decay and a further b-jet do not alter the E_T^{miss} shape, the $Zb\bar{b}$ has a well separated median and accumulates to lower energies). This can be ascribed to the softer spectrum of neutrinos coming from bottom quark decays. The feature will turn particularly useful when background will be directly searched in data, because it allows to significantly separate one source from another. Here the optimal threshold has been identified between 30 and 35 GeV.

When the cut is applied, signal efficiency reduces by 0.817 ± 0.014, $t\bar{t}$ by 0.837 ± 0.002 and $Zb\bar{b}$ by 0.390 ± 0.002, hence effectively increasing the signal to background ratio.

Processes having di-boson producing neutrinos (as WZ and WW) and semileptonic tau decays (as in ZZ), though may extend up to high energies, have a smaller cross sections, thus a negligible impact.

Constraints on the W mass

To reproduce the lepton and neutrino produced by a W boson coming from a top, muon and electrons are sorted between objects whose selection has been earlier optimized. Leptons found too close to b-jets and light-jets candidates ($\Delta R < 0.3$), or having been exploited to build the Z mass are removed from the candidates. On these leptons with enhanced purity the cuts $p_T(e) > 20\,\mathrm{GeV}/c$ and $p_T(\mu) > 15\,\mathrm{GeV}/c$ are imposed. These thresholds have demonstrated to deliver a very good purity, a high selection efficiency and the best possible $t\bar{t}$ and $Zb\bar{b}$ rejection.

Since the longitudinal momentum of the neutrino cannot be directly measured, the W boson can be reconstructed only in the transverse plane. Therefore, a wide transverse mass distribution will replace a narrow peak around the nominal W mass value, but it will prove sufficient to give confidence in a good gauge boson identification.

Given the component $(p_x(l), p_y(l))$ of the lepton $p_T(l)$ and reconstructed transverse component of the neutrino momentum, the transverse mass of the lepton combination is defined from the equation:

$$M(l\text{-}E_T^{miss})_T = \sqrt{\left(p_T(l) + E_T^{miss}\right)^2 - \left(p_x(l) + p_x(\nu)\right)^2 - \left(p_y(l) + p_y(\nu)\right)^2}, \tag{3.4}$$

that demonstrates to be equal to

$$M(l\text{-}E_T^{miss})_T = \sqrt{2p_T(l)E_T^{miss}\,(1 - \cos\phi)}, \tag{3.5}$$

where ϕ is the azimuthal angle between the lepton and neutrino.

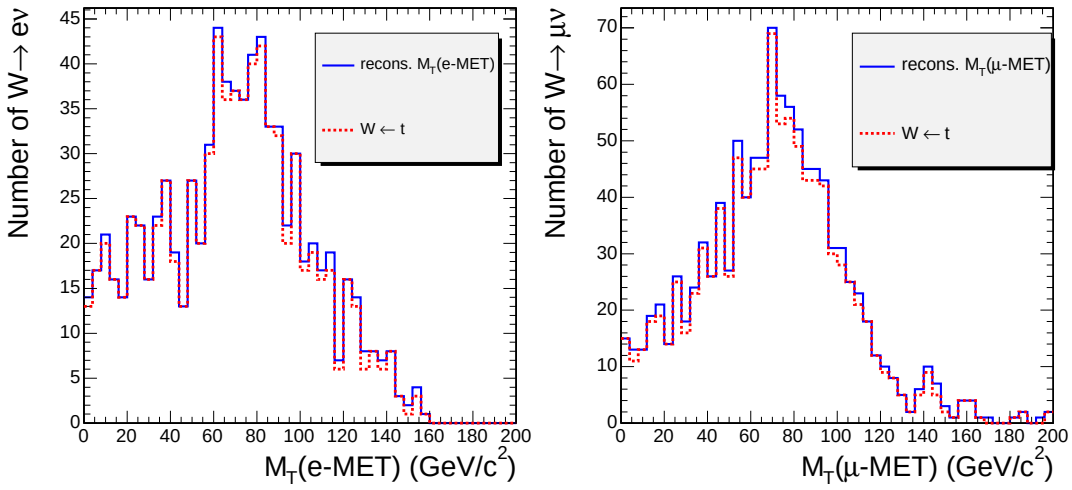

Figure 3.33: The W boson after requiring a l^+l^- consistent with a Z, reconstructed with a third electron (left) or muon (right) and large missing energy. The approximative Jacobian shapes and the agreement with the generated objects ensure that the boson has been properly identified.

Distributions computed from Eq. 3.4, with the additional requirement to keep the lepton-missing transverse energy combination with the invariant mass closest to the mass of the $W^{\pm}$, are showed in Fig. 3.3.3 for the $W \to e$ and $W \to \mu$ cases. The 'shoulder' at low energy and the steep fall after the W mass indicate that a Jacobian shape, smeared by the finite detector energy resolution, is correctly reproduced.

A problem this reconstruction could suffer is a mis-identification with leptons from Z, that are produced in pair with a similar energy and in all possible directions. The comparison with the distribution of a 'matched' transverse mass, in which a reconstructed lepton closer than $\Delta R = 0.3$ to a generated lepton coming from W is used, shows that the agreement with true object is good in all the mass range, for both muons and electrons.

Any event containing more than one good W candidate is rejected.

The W boson transverse mass reconstructed in this way is plotted in Fig. 3.34, where it is compared with the matched object and the full background.

The lower logarithmic plot confirms that the most important contribution comes from $Zb\bar{b}$, $t\bar{t}$ and partially Z+jets, stacked on top of the multi-leptonic ZW, ZZ processes that are much smaller.

As is clear from Eq. 3.3.3, a significant high transverse mass tail in the distribution may be originated from uncorrelated hard leptons or neutrinos. This happens when more than one leptonic decay (as in $t\bar{t}$ background) or high-flavour semileptonic decay, together with high p_T leptons (as in $Zb\bar{b}$) enter the game. A possible upper cut in the $M_T(l\text{-}E_T^{miss})$ is optimized as in the Z selection: the reconstructed mass distribution is used to scale the signal and background number, then the quantity s_x (that is proportional to the $N_S/\sqrt{N_B}$ ratio) is calculated for different cuts. The results for the significance of signal against $Zb\bar{b}$ and

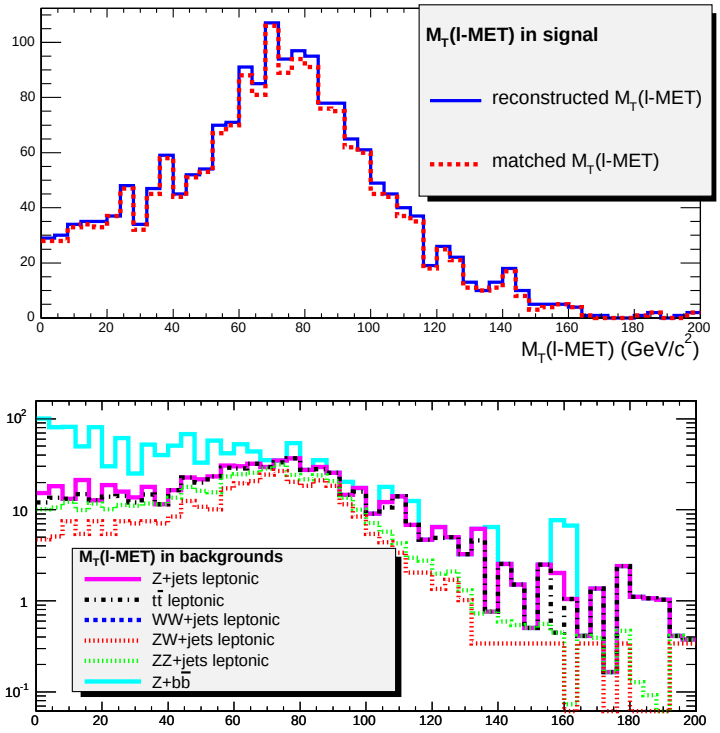

Figure 3.34: The distribution of the transverse invariant mass of the $e^{\pm}$ or $\mu^{\pm}$-missing energy combination in the signal sample and in the different background samples. The background distributions have been normalized assuming an integrated luminosity of L=10fb^{-1}.

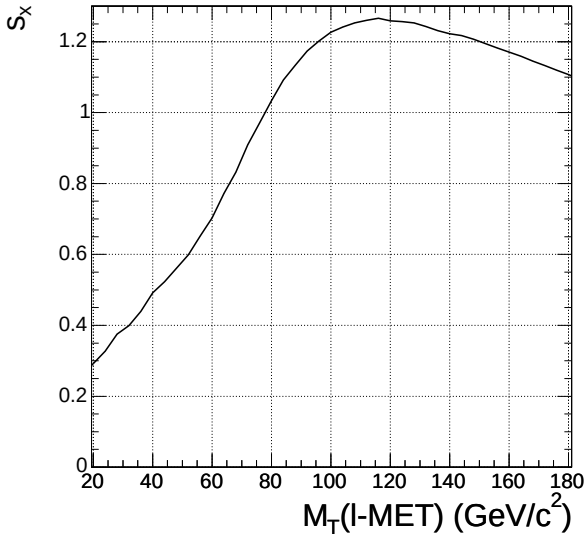

Figure 3.35: The curve for the optimization of the upper cut in the W transverse invariant mass. The significance is maximized when $M_T(l\text{-}E_T^{miss})$ <120 GeV/c^2 is chosen, and it is quite robust in variation around this threshold – due for instance to systematic effects.

$t\bar{t}$ are displayed in Fig. 3.35. The optimal mass cut is found between 115 and 120 GeV/c^2. The displayed curve is quite smooth around this best point: from it, one can guess that the final sensitivity of the experiment would be relatively lowered by 5%, only if resolution on transverse W mass is worst than about 20 GeV/c^2.

If these results are compared with efficiencies in Tab. 3.15, the request of the third lepton appears as a major tool in rejecting the more pernicious background.

In order to reduce the statistical error on efficiency estimates, a control sample has been exploited, where the statistic has been artificially increased by opening the cut in the Z mass, toward $\Delta M = 100\,\mathrm{GeV}/c^2$. Relative efficiencies are then computed on this control sample. This seem reasonable for the $t\bar{t}$ background, since enlarging the Z mass window has impact only in the amount of surviving events. The relative efficiencies of cascade cuts are insensitive to it, because kinematical properties of $t\bar{t}$ are almost uniform even in a wide range on the $M(ll)$ spectrum.

3.3.4 Light-jets and b-jets specific analysis

General performance of b- and light jets reconstruction are already discussed above. Below, a study for an optimization oriented to maximize the background reduction is presented.

Selection of the final state b-jets

The jets with beauty are selected with the criteria optimized in Sec. 3.2.3 and applied to the signal and background processes. The condition for the presence of both a reconstructed Z

Table 3.16: The efficiencies for the reconstruction of a Z boson plus a W, both in the leptonic decay channels, from all the considered background. Kinematic cuts discussed in the test are applied to leptons and invariant mass. All values are in percent.

Selection	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Zb\bar{b}$	$ZZ \to 4l$
rec. $W \to e\nu$ (no $M_T(l\text{-}E_T^{miss})$ cut)	10.6 ± 0.4	0.0133 ± 0.0012	0.68 ± 0.06	0.766 ± 0.013
rec. $W \to \mu\nu$ (no $M_T(l\text{-}E_T^{miss})$ cut)	12.0 ± 0.4	0.0133 ± 0.0012	0.68 ± 0.06	0.766 ± 0.013
one good $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	20.7 ± 0.5	0.0130 ± 0.0012	0.120 ± 0.011	4.19 ± 0.07
Selection	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
rec. $W \to e\nu$ (no $M_T(l\text{-}E_T^{miss})$ cut)	0.021 ± 0.004	0.131 ± 0.019	0.295 ± 0.014	0.007 ± 0.003
rec. $W \to \mu\nu$ (no $M_T(l\text{-}E_T^{miss})$ cut)	0.021 ± 0.004	0.131 ± 0.020	0.352 ± 0.017	$\sim 3 \cdot 10^{-5}$
one good $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	0.013 ± 0.003	0.27 ± 0.04	1.86 ± 0.059	0.005 ± 0.002

and W is inserted before this selection. The distribution in transverse momentum for all the sources is displayed in Fig. 3.36(left), where the rescaling for the cross section is applied and the signal distribution (upper) is compared with the tagged jets matched with b quarks.

The backgrounds entering this selection are only those with energetic b-jets, namely the $Zb\bar{b}$ and top–anti-top production, that results in a pair of b-jets for each event. All the other sources appear nicely suppressed from the multi-lepton request. Even the Z+jets process, that features a large cross section and has objects that can potentially be mistagged, appears as un-sizable histogram over the $t\bar{t}$ production.

When the comparison of signal is restricted to the $Zb\bar{b}$ and $t\bar{t}$ sources and events are normalized to the same number, a plot like Fig. 3.36(right) can be obtained. While the b-jets from the top quarks in the signal and in the background share exactly the same kinematics and they are not distinguishable at this level, bottom quark production at the primary vertex demonstrates to have a softer spectrum. The optimal p_T threshold turns out to be in the $45 \div 50\,\mathrm{GeV}/c$ range, which coincides with the requirements for a maximal tagging purity and efficiency. Thus the $50\,\mathrm{GeV}/c$ cut confirms to have the better property, both for efficiency and background rejection.

Two further variables that are relevant in the b-jets identification, as the discriminator and the number of tagged jets, are estimated for signal and major background after the $W+Z$ request and showed in Fig. 3.37, where the same number of events is compared. The shape of discriminator distribution for the signal is shared with the $t\bar{t}$ and $Zb\bar{b}$ processes. The two-fold

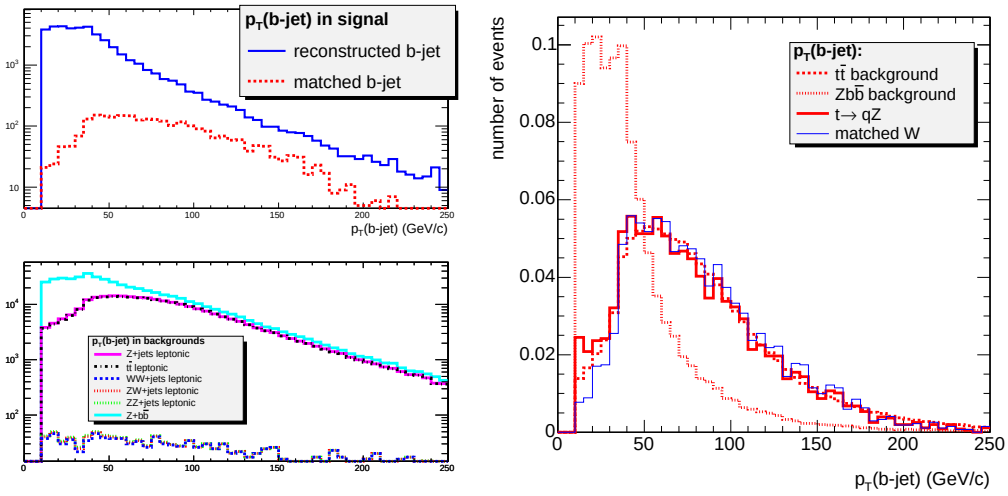

Figure 3.36: Left: The comparison of $p_T(b\text{-jet})$ between signal (upper) and background (lower). In the signal, b-tagging demonstrates to approach the maximum purity only for high p_T; the background distribution, where histograms are plotted stacked, shows that only $Zb\bar{b}$ and $t\bar{t}$ are sizable. Right: The same distribution for signal and relevant background, when normalized to the same number of events.

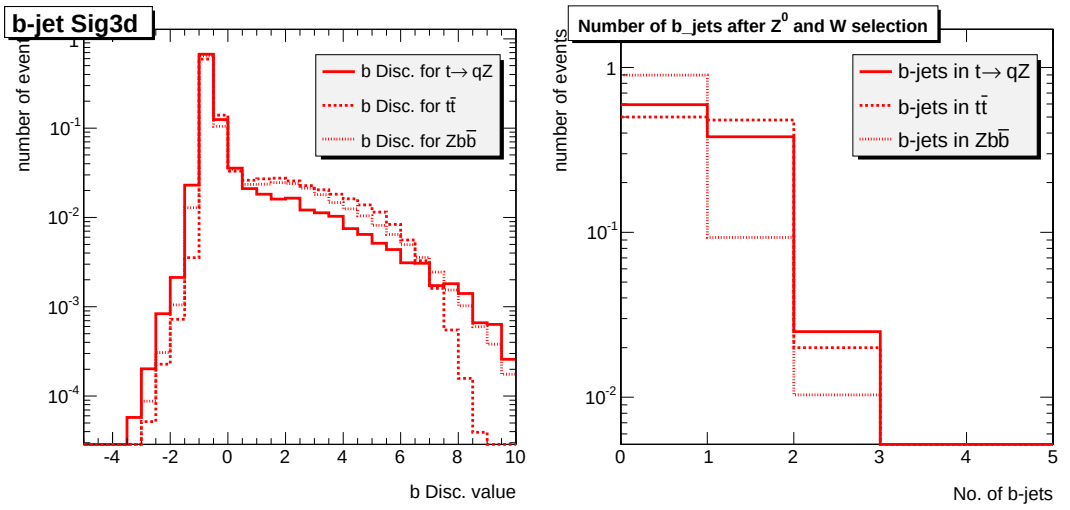

Figure 3.37: Two further variables are inquired to look for a separation between signal and background, as the discriminator distribution (left) and the number of b-tagged jets after the tagging procedure implemented in the reconstruction (right). Feasibility of the discrimination based on these variables is discussed in the text.

nature of b-jet multiplicity in these states is reflected in a significant higher $Disc$ value in the region where a jet is supposed to contain a b quark. When a very high discriminator is asked, the jets in $t\bar{t} \to 2l$ background have a probability to contain one or more b-jet that is lower than the $Zb\bar{b}$ events. In fact, when selecting the same charge and opposite flavour leptons in the $M(ll)$ distributions, in the $t\bar{t}$ background leptons from b decay may be likely associate to lepton from W. The possibility to mis-identify leptons coming from both b-jets as leptons from the gauge boson is larger in the $t\bar{t}$ than in $Zb\bar{b}$: as a consequence, the high-likelihood tail in the discriminator turns to be larger for the latter process.

While the discriminator confirms to be powerful in measuring the probability for a jet to be tagged, no further cuts (lower or even upper) can be devised to deplete the background, so the $Disc > 2$ selection is not modified.

Percentages of event with 0, 1, 2 or more b-tagged jets is plotted in Fig. 3.37(right). If only the events with a single b-jet are accepted, the histograms shows how large is the fraction of surviving background. Almost 90% of b-jets from $Zb\bar{b}$ does not success in the pass the stringent selection, while another 1% of double-tagged events is eliminated. Half of $t\bar{t}$ sample passes the cut, confirming as the most resistant background; due to the hard multi-lepton request, only 2% of these events have two tagged b-jets. The efficiency of a single b-jet for the signal is around 38%, with a 2% of mistagged jets that are rejected. In summary, the single-tagging requirement has not a major role in rejecting these backgrounds, but testifies that processes different from these are very unlike to pass this b-tagging. Standard Model processes with heavy bosons and a single b (as single-top production in the t-channel or $Z + b + j$) could be reduced by some $\sim 9\% \times 9\% \simeq 0.8\%$, thus confirming their non–relevance in the selection cuts optimization. Table 3.17 (upper rows) indicates the cumulative efficiency after the selection of a single b-jet, and is a different way to express the suppression of all the background without a multi-b-jet production.

As underlined earlier, here some of the background sources are eagerly suppressed and a '0' appears in the table for the first time. It corresponds to the fact that no simulated events have passed the whole selections, so only an upper limit can be quoted. These upper limits are clearly related to the size of simulated event samples that, for not too relevant background, is keep limited to save computing time. Now, if the count of the number of background event is assumed to be Poissonian, the observation of no events corresponds to an upper limit of the confidence interval of the Poissonian mean equal to 2.3 at 90% C.L. and 3.0 at 95% C.L.. When the most conservative limit is assumed and the number of generated events is kept in mind, these zeroes can be translated in $< 2 \cdot 10^{-5}$ for the $Z + j_{85-150}$, $< 1 \cdot 10^{-5}$ for the $Z + j_{150-300}$, $< 3 \cdot 10^{-5}$ for the $WZ \to 3l$ and $< 3 \cdot 10^{-5}$ for the $WW \to 2l$.

The results in Tab. 3.17 are directly linked to the performance of b-tagging algorithm, whose efficiency and purity have been optimized earlier. To reproduce the tagging efficiency by counting in these samples, the ratio between percentages in Tab. 3.17 and in Tab. 3.16 is the starting point. It has to be scaled by the b-jet multiplicity in each sample, than the

Table 3.17: The efficiencies for all the considered background when the presence of one b-tagged jet is added to the previous selection (upper). In the last row, an estimated of the absolute b-tagging efficiency drawn from comparison of this cut with the previous, and corrected for the portion of p_T(b-jet) samples, is reported.

Selection	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Zb\bar{b}$	$ZZ \to 4l$
one good Z + one good W + one b-jet	6.09 ± 0.31	0.0044 ± 0.0003	0.079 ± 0.002	$\sim 3 \cdot 10^{-5}$
Selection	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one good Z + one good W + one b-jet	0	0	0	0
	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Zb\bar{b}$	$ZZ \to 4l$
b-jet extrap. eff.	0.375 ± 0.030	0.405 ± 0.024	0.308 ± 0.027	–

fact that only a portion of the p_T spectrum has been considered is to taken into account. The b-tagging efficiencies estimated in this way are reported in Tab. 3.17 (bottom row), and can be a hint for an evaluation of tagging power from real data, thus avoiding to measure performance on simulated events. These values have to be compared with the efficiency study that has been addressed when optimizing the tagging. To accomplish that task, there the reconstructed distribution were compared with the one from the generated ones, and studied as a function of momentum and discriminator threshold. When that result are re-considered and the threshold $Disc = 2$ is chosen, one obtains that the efficiency integrated in the whole spectrum reads 0.496 ± 0.018, remarkably different from the value for signal quoted in the table above. The discrepancies could be ascribed to the fact that the sample of b-jets that has selected here to estimate the efficiency is not so pure, *i.e.* they not all come from the top decay. If the value in last row of Tab. 3.17 is instead rescaled for the b-jet purity integrated over all p_T (see for instance Fig. 3.6(right)), around 75%, a value in a nice agreement with that efficiency is obtained.

At this level, one may like to inquire how much the selection included in the set of devised cuts are correlated. This is important not only to manage the total error on the sample counting, but even to see if each optimized cut can be applied independently from the others that, as it will be shown at the end of the chapter, turns useful to infer *a posteriori* considerations on other possible background sources.

While the three-leptons selection can reasonably assumed to be the product of the three separated selection, one could suspect a relation between the number of b-tagged jets and the leptonic requirements. In particular, a cut in the missing transverse energy should reduce the number of b-jet surviving the tagging algorithms, since the b quark are bound to neutrinos

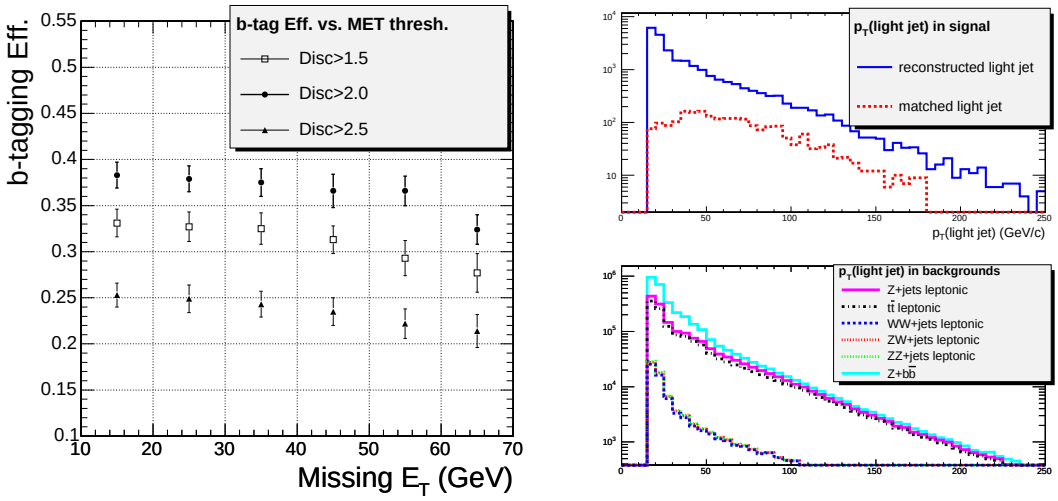

Figure 3.38: Left: Tagging efficiency for b jets coming from a top, when different lower limits for the missing transverse energy from the W are chosen. Three different discriminator thresholds are also used. Right: The distribution of the transverse momentum of light-quark jet candidates in the signal sample and in the different background samples. The corresponding distribution for jets matched to the generator-level light quarks from the FCNC top decay is also shown. As usual, background histograms are stacked and show how $t\bar{t}$, $Zb\bar{b}$ and Z+jets stay on top of the di-boson background, whose jets are not from a hard production, thus they have no impact anymore.

by their semileptonic decays.

In the Fig. 3.38(left), the analysis is repeated for six different E_T^{miss} thresholds and the efficiency is evaluated each time. As just a comparison between efficiencies is interesting here, only the 'raw' estimate is adopted; three reference values for the discriminator thresholds are also adopted.

Some observations are here straightforward:

- in all the three samples, the raw tagging efficiency seems slightly decrease, thus motivating the intuition;
- in all the three samples, these variation is quite well included in the error bars (only statistical error are considered here). A possible correlation coefficient turns out to be about 0.4, thus a correlation term has not significance on top of the total error estimate;
- by the way, the result that efficiency is optimal with a lower $Disc$ between 1.5 and 2.5 is offered also from this check.

In other words, the portion of selected missing energy seems well separated by the lower missing energy produced by neutrino in $b \to ql\nu$. Therefore, independences of all cuts is fully preserved.

Selection of the final state light-jets

Transverse momentum distribution of light-jets in the $t \to qZ$ analysis is shown in Fig. 3.38(right). These distributions include all jets reconstructed in the event, once cleaned from the selected b-jets and 'ghost' clusters in the calorimeter. In the upper plot the reconstructed light-jets transverse momentum is compared with objects matching a u or c quark. The loss in purity is well visible here: as explained below, it can be ascribed to a mis-identification with b-jets and motivated by the need to reject the dsg contamination and equalize the response between u and c quarks.

Since the cross sections for the di-boson production followed by leptonic decays are rather small, the distribution from the background (lower plot, stacked histograms) is dominated by the $t\bar{t} \to 2l$ events, where at least one of the b-jet is mis-interpreted as coming from a light quark. The same effect allows also the important $Zb\bar{b}$ to pass also this selection. As the great part of light-jets accumulates below some $\sim 40\,\mathrm{GeV}/c$, only a tail will be selected with the p_T $>60\,\mathrm{GeV}/c$ cut, and it is not possible to found a different optimal cut. On the other hand, from that energy all these backgrounds nicely decrease with an exponential law: therefore, the pre-selection in p_T applied in the generation process is well motivated.

Cross-checking this plot with the Fig. 3.36(left) is another view to guess which background sources turn out to be important. If both the light and the b-jet are required in the event, background distribution roughly factorizes, killing most of the processes but $t\bar{t}$, $Zb\bar{b}$ and Z+jets. When the presence of the three hard leptons is demanded, the latter source is highly suppressed, while the others will be still there and challenge the reconstruction of the signal.

3.3.5 Constraints on the top with SM decay

The reconstruction of the top quark following the standard decay is an important step in both the analyses, because it constraints the b-jet and the W boson in a final state to fill an invariant mass window, thus further limiting the contribution of processes that do not have a top.

Since only W leptonic decays are involved, the full reconstruction of the top decay product is not possible without additional kinematic constraints. To avoid the introduction of the biasing effects associated to these constraints, no full mass reconstruction is performed here and only the 'transverse mass' is evaluated. Its definition closely resembles that for W transverse mass:

$$M_T(bW) = \sqrt{\left(E_T(b) + p_T(l) + p_T(\nu)\right)^2 - \left(p_x(b) + p_x(l) + p_x(\nu)\right)^2 - \left(p_y(b) + p_y(l) + p_y(\nu)\right)^2},$$

where the transverse momenta for the lepton $(p_x(l), p_y(l))$ and for neutrino $(p_x(\nu), p_y(\nu))$ are explicited, and $E_T(b\text{-}jet)$ is assumed to measure the transverse energy of the originary b

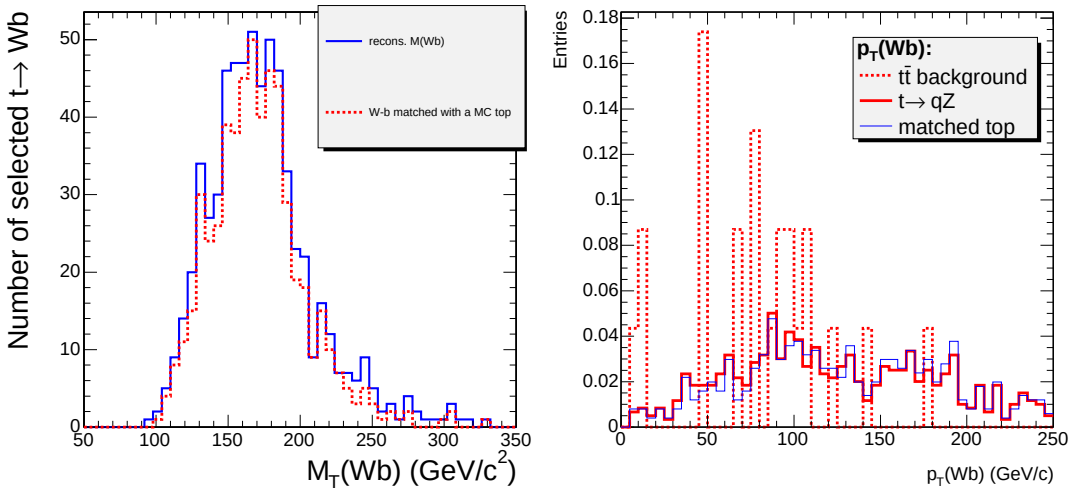

Figure 3.39: Left: The reconstructed transverse invariant mass of the top following the standard decay, along with the same distribution made with 'matched' W and tagged jets close to the b quark. Right: transverse momentum of the reconstructed top, from signal, matched signal and the few events surviving from the $t\bar{t}$ background.

quark.

Exploiting one of the b-jet and W boson whose selection has been optimized above, the distribution depicted in Fig. 3.39(left) is obtained. A reference distribution is over-imposed, that corresponds to the M_T of a lepton from W close to a generated lepton ($\Delta R < 0.3$) together with a jet close to a generated b quark. The shape expected for a transverse mass is well reproduced and the match in the $100 \div 250\,\mathrm{GeV}/c^2$ range is satisfied within few percent level, confirming the correctness of the reconstructed object.

At this level, the background to the $t \to qZ$ has been effectively reduced to the top pair and the $Zb\bar{b}$, with a small amount of events surviving the three leptons and b quark requirements. Distribution of the reconstructed top transverse momentum is almost uniform, as shown in Fig. 3.39(right). The few $t\bar{t}$ events satisfying the reconstruction criteria seem to follow the same distribution, thus no p_T cut on the top is possible.

A full comparison with all the background sources is offered by Fig. 3.40(left). The fact that all optimization studies for this signal involve only the $Zb\bar{b}$ and $t\bar{t}$ founds here a clear motivation. In each event, only one such object selected in this way is asked for.

In order to improve the statistical properties of this transverse mass region, the distribution tails can be cut away. If the significance s_x is plotted as a function of the upper and lower mass threshold, the plot in Fig. 3.40(right) presents. While no lower cuts are indicated by the function on the left, the selection is slightly improved if one requires $M_T(bW) < 220\,\mathrm{GeV}/c^2$. Robustness around this cut is again satisfactory: it can be evaluated that, if the threshold value is shifted by about 10%, the decrease of s_x induces a relative variation in the signal efficiency around 2%.

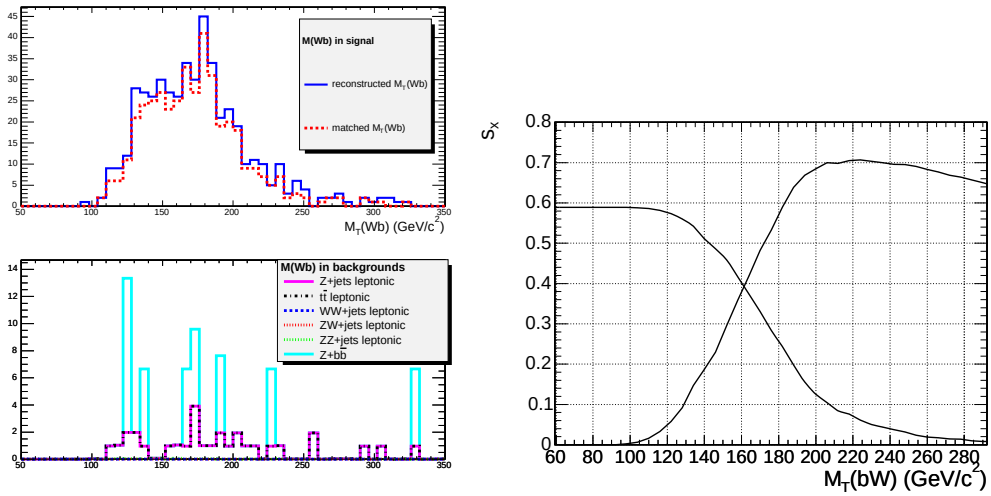

Figure 3.40: Left: The invariant mass of the e^+/e^- or μ^+/μ^- and missing energy and b-jet combinations for signal and various backgrounds. The background distributions have been normalized assuming an integrated luminosity of L=10fb^{-1}, and it shows that the only events that accumulate in the top mass window come from $Zb\bar{b}$ and $t\bar{t} \to 2l + X$. Right: The efficiencies for the reconstruction of the invariant mass for the top decaying $t \to Wb$, selected with the criteria specified in the text. Only the $t\bar{t}$ and $Zb\bar{b}$ are survived at this level. Efficiencies on the background have been evaluated by enlarging the Z mass window, so limiting the statistical error.

The resulting cumulative efficiency for the signal and most important background are listed in Tab. 3.18, with all values in percent. At this level, even the ZZ process has no more events in the selected samples, and its upper limit is $2 \cdot 10^{-5}$. As anticipated at the beginning and extensively assumed, in these last cuts only $t\bar{t}$ and $Zb\bar{b}$ dominate the scene.

It is important to observe that, since almost of background events numbers are approaching to zero, here optimization procedure is close to its limit. While high values for the statistical significance are successfully reached, further heavy cuts would result in completely suppress the background, thus making the experiment unpredictive. This limit is clearly due to the amount of statistic generated, that in turn is related to the assumed integrated luminosity. In the following chapter, the surviving background will be estimated in a way that avoids the wide statistical fluctuations for small numbers, and can be extended to larger integrated luminosities.

3.3.6 Constraints on the top with FCNC decay

After the products of the SM top decay have been identified, the final step of the analysis process is to combine the Z^0 candidate with good light-quark jet candidates. The combination whose invariant mass lies closest to the nominal top mass is assumed to be the correct one.

The result shown in Fig. 3.41(upper) along with the 'matched' distribution, that is the one for a top quark made with a reconstructed jet close to a u/c quark and a matched Z, only

Table 3.18: The efficiencies for the reconstruction of the invariant mass for the top decaying $t \to Wb$, selected with the criteria specified in the text. Only the $t\bar{t}$ and $Zb\bar{b}$ have survived at this level. Efficiencies on the background have been evaluated by enlarging the Z mass window, so limiting the statistical error.

Selection	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Zb\bar{b}$	$ZZ \to 4l$
one good Z + one $t \to Wb$	5.26 ± 0.11	$(2.37 \pm 0.28) \cdot 10^{-3}$	$(3.4 \pm 1.3) \cdot 10^{-2}$	0
Selection	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one good Z + one $t \to Wb$	0	0	0	0

when a W is reconstructed and a b-jet is close to a b quark. The comparison demonstrates that, while a narrow peak accumulates around the right top mass, a significant excess of event different from a genuine top is present. This 'combinatoric' background comes from having collected in the same point of the phase space the Z and the light-jets found around, and it will be successfully identified by a fit procedure.

The signal distribution obtained in the last selection is here adapted to an analytic function, in order to disentangle the combinatorial background to the genuine signal. Here the type of the function $f_{B+S}(x)$ adopted is

$$f_{B+S}(x) = \frac{S_0}{\sqrt{2\pi}\sigma} e^{-\frac{(x-M_0)^2}{2\sigma^2}} + f_B(x), \tag{3.6}$$

i.e. a normalized Gaussian with an additional function parameterizing the background. S_0, σ and M_0 are parameters to be found, with the latter expected to reproduce the top mass. An empirical function delivering a quite good convergence has found to be composed from a constant and a function that rises linearly from a point x_0, then it bends and fall down for $x \gg x_0$, with an exponential constant M_1. The functional form is:

$$f_B(x) = B_0 \left(1 + (x - x_0) e^{-M_1 (x - x_0)}\right), \tag{3.7}$$

where the further parameter B_0, x_0 and M_1 must be determined by the fit. The composed function has been applied to the distribution in the large (x_L, x_H)=(100,300) GeV/c^2 range, in order to maximize the statistic power of the calculation. The result of this operation is presented in Fig. 3.41(lower), where the signal has the error bars obtained from a quadratic sum of Poissonian uncertainties, and both the total and the background function are overimposed. The fit converges toward a mass value that accommodates the input top mass ($m_t = 175\,\mathrm{GeV}/c^2$) and resolution on the central value (with only statistic error included) is better than 1%. The number of signal and combinatorial background can be found by an analytic integration of the fitted function. For the signal, it results:

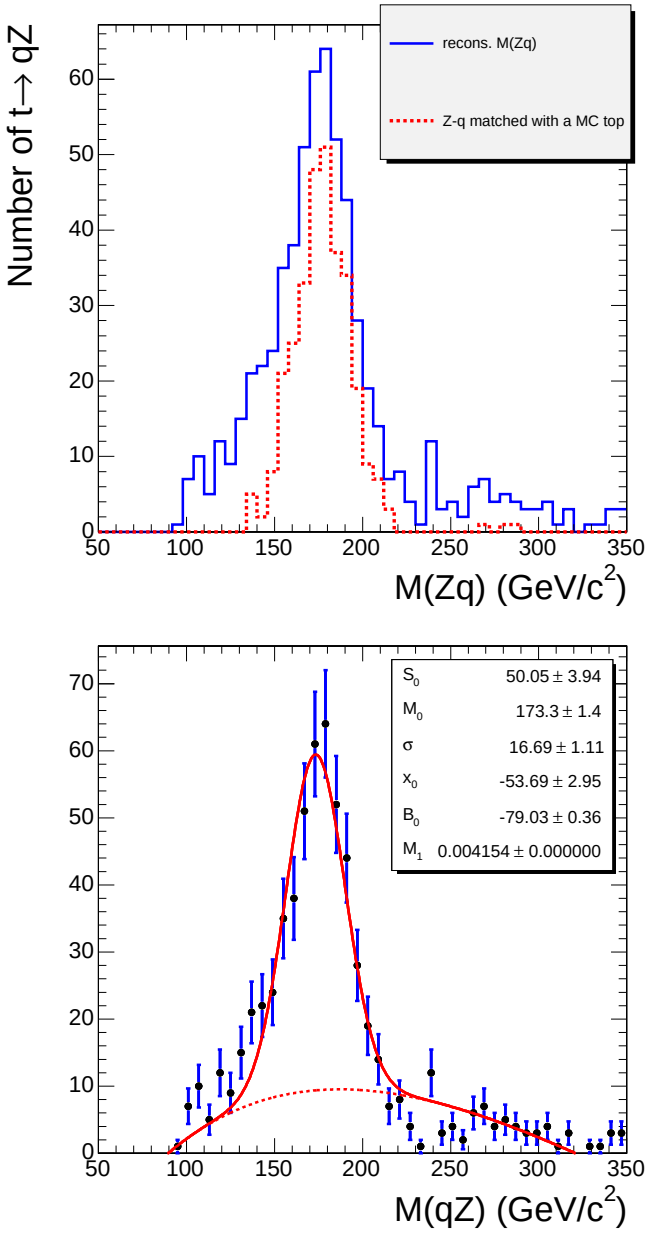

Figure 3.41: Upper: The invariant mass $M(qZ)$ of the top quark decaying with the anomalous process, obtained with the selected Z and selected light-jet. A bias of combinatorial events stays under the real top signal. Lower: The result of the fit on the top FCNC signal distribution with a Gaussian (for the 'genuine' signal) summed to a linear plus exponential (for the combinatorial background) function. Total function (solid) and background (dotted) are over-imposed and resulting fitted parameters are in the table.

$$S = \frac{S_0}{\sqrt{2\pi}\sigma}\int_{x_L}^{x_H} e^{-\frac{(x-M_0)^2}{2\sigma^2}}dx = \tag{3.8}$$
$$\frac{S_0}{\sigma}\int_{\frac{x_L-M_0}{\sqrt{2}\sigma}}^{0} e^{-y^2}dy + \frac{S_0}{\sigma}\int_{0}^{\frac{x_H-xM_0}{\sqrt{2}\sigma}} e^{-y^2}dy =$$
$$S_0\sqrt{2\pi}\sigma\left(\frac{1}{2} - Erf\left(\frac{M_0 - x_L}{\sqrt{2}\sigma}\right)\right) + S_0\sqrt{2\pi}\sigma Erf\left(\frac{x_H - M_0}{\sqrt{2}\sigma}\right),$$

where $Erf(y)$ is the Error Function. Since the integration range has been choose with $x_L \ll M_0$ and $x_H \gg M_0$, the first Erf converges to 0.5 and the second to 1: therefore, $S_0\sqrt{2\pi}\sigma$ is a very good approximation ($\mathcal{O}(10^{-3})$) of the integral under the mass distribution. Once it has been rescaled for the histogram binning, that is 0.167 events/ GeV , the fitted number of signal events turns out to be $N_S = 349 \pm 36$, that corresponds to an efficiency $\epsilon_S = (4.4 \pm 0.4)\%$. This is assumed as the final estimate of the selection efficiency for the $t \to qZ$ signal.

For the combinatorial background, it results:

$$B = B_0\left(x_H - x_L + \frac{1}{M_1}\int_{\frac{x_L-x_0}{M_0}}^{\frac{x_H-x_0}{M_0}} ye^{-y}dy\right) \tag{3.9}$$

that can be integrated analytically.

As a consequence, the fitted Gaussian function can be considered as an effective line-shape of the top quark decaying with the anomalous coupling. It can be compared with the 'matched' invariant mass distribution, that is performed using a light-jet closer ($\Delta R < 0.3$) to the generated one, and the matching condition for the Z and the Standard Model top is ensured. This comparison results in Fig. 3.42(left), that actually demonstrates the good quality of this final reconstruction.

Once the selection procedure is repeated for the all backgrounds, only a tiny number of events enters the mass window between 100 and 250 GeV/c^2 – that as demonstrated is the only place where a real top quark exists. At this level, this number is so little that the vertical scale in the lower plot of Fig. 3.42(right) has a fractional number of events. Far to be unphysical, it has only to be interpreted as the number of events surviving for each 10 fb^{-1} of integrated luminosity once the set of cuts has been applied. The predictivity on efficiencies dictated by the size of simulated event samples is here pushed to its limit: that is why estimation of efficiencies is performed by widening the Z mass window, as detailed above – nevertheless statistic error is still large.

By a parton-level studies it could be shown that the majority of the $t\bar{t}$ pairs are produced in a "back-to-back" configuration in the transverse plane. This behaviour, being characteristic of the top pair production, could be a handle to reject the background coming from other sources,

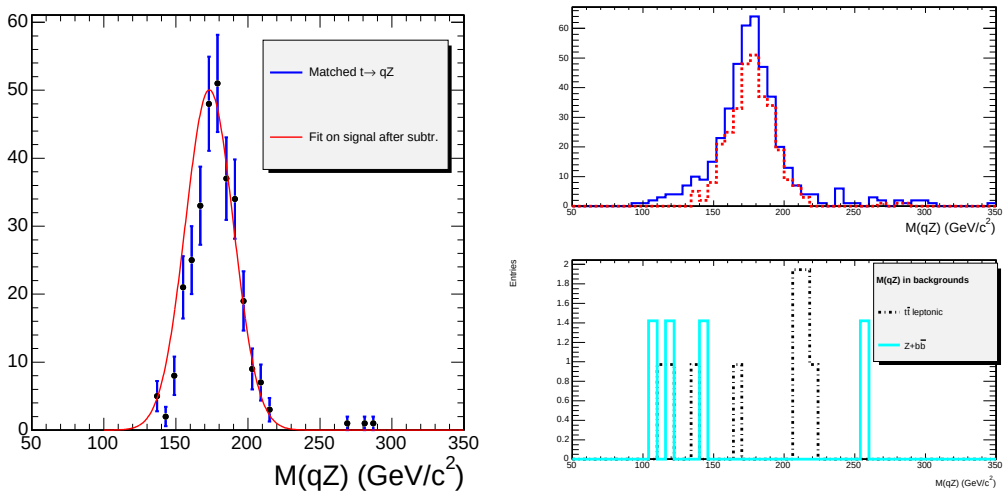

Figure 3.42: Left: The result of the Gaussian part of fit on the top FCNC signal distribution, compared with the histogram of the matched object. The agreement is good and demonstrates that the aim to extract the lineshape of the top quark from its decay, on top of a rich background has been reached. Right: The effect of the search for the FCNC decay in the $M(qZ)$ invariant mass plot, as resulting when applied to the signal and background. Only the two backgrounds that are supposed to deliver a sizable contribution are indicated; number of count is scaled to the rate foreseen after $10\,\mathrm{fb}^{-1}$.

as the $Zb\bar{b}$ or combinatorial under the signal. This was done in Ref. [96] where, adopting a set of slightly different cuts, it demonstrated useful to reject some fake top candidates. On the other hand, the angular distribution of top quark is known to be one of the variables most suffering in going from the LO to next orders: in Ref. [22], for instance, it was shown that azimuthal correlation may vary by $\sim 50\%$ or more when going from Monte Carlo (HERWIG) to NLO calculations. Though here the amount of background will be evaluated by counting in a specific control region, it will not include any angular cut, so this uncertainty would be an important source of systematic effects. In the present analysis, this price would be payed against a very modest reduction of the background – that is already very little. Therefore, here it is consider safer to avoid apply any angular selection, and no further cut is devised. The combinatoric background, though sizable, it is readily identified by the fit procedure on the signal, that has been explained. In this way, instead of insisting on a stronger reduction of the background (that with a so small amount of event does not imply any enhancement of the statistical properties), the focus is put on the signal distribution, by cleaning it from accidental objects entering the selection window.

The efficiencies for the final selection of the FCNC channel are contained in Tab. 3.19, for the signal and all the simulated background sources.

Table 3.19: The efficiencies for the final selection of the FCNC channel, *i.e.* the presence of a light-jet together with the objects reconstructed above (first row) and a further selection in the 100-250 GeV/c^2 mass window (last row). All the backgrounds but $t\bar{t}$ and $Zb\bar{b}$ are suppressed up to the level explained in the text. Efficiencies are quoted in percent.

Selection	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Zb\bar{b}$	$ZZ \to 4l$
one Z + light-jet + $t \to bW$	5.50 ± 0.26	$(9.91 \pm 1.4) \cdot 10^{-4}$	$(1.02 \pm 0.32) \cdot 10^{-3}$	0
100< $M(qZ)$ <250	5.32 ± 0.24	$(9.91 \pm 1.4) \cdot 10^{-4}$	$(0.76 \pm 0.24) \cdot 10^{-3}$	0
Selection	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one Z + light-jet + $t \to bW$	0	0	0	0
100< $M(qZ)$ <250	0	0	0	0

3.3.7 Summary of the $t \to qZ^0$ analysis

An efficient procedure has been devised here to detect the $t\bar{t} \to (Zq)(W\bar{b}) \to (l\nu b)(llq)$, on top of a wide variety of background. The signal selection efficiency has turned out to be of the order of percent, while the surviving background event number is of the order of unity at 10 fb^{-1}.

The main selection requirements for this channel are summarized briefly below; each item corresponds to a row in Table 3.20:

1. since the presence of a Z boson is a powerful distinctive signature of the signal, the events are triggered by the standard "double electron or double muon" stream, both at L1 and HLT;

2. well isolated electron and muon are sorted with a p_T greater than 15 GeV/c for muons and 20 GeV/c for electrons, and combined in an invariant mass plot. The event must have one only Z in a 10 GeV/c^2 -wide window around its nominal mass;

3. a third isolated lepton, featuring the same quality cuts is asked for. The combination of this lepton with the missing energy greater than 35 GeV is imposed to have a transverse mass less than 120 GeV/c^2 . Events with more than one good W candidate are rejected;

4. candidate events are asked for a single b-tagged jet (b-tagging discriminator greater than 2), having $p_T(b\text{-jet}) > 60$ GeV/c;

5. the invariant mass of the $W^{\pm}$ and the b-jet must have a tranverse mass below 220 GeV/c^2 ;

6. each event must contain at least one jet which is incompatible with coming from a b quark and which satisfies $p_T > 60$ GeV/c. The combination of this object with the Z^0 candidate must have an invariant mass in the range 100< $M(qZ)$ <250 GeV/c^2 .

Each cut has been tuned with the aim to maximize the statistical significance of the result that, as it will be proved in the next chapter, is proportional to the $N_S/\sqrt{N_B}$ ratio.

Table 3.20: The efficiencies (in percent) for the most important selections cuts in the $t \to qZ^0$ analysis. Zero values mean that no events from the chosen event sample is expected to survive; statistical upper limits are detailed in the text.

Selection	Signal $t \to qZ$	$t\bar{t} \to 2l$	$Z(\to ll)b\bar{b}$	$ZZ \to 4l$
HL Trigger	89.0 ± 1.3	42.95 ± 0.07	92.3 ± 0.3	59.1 ± 0.2
one good $Z^0 \to ll$, after $\Delta M(ll)$	42.8 ± 0.7	1.500 ± 0.013	34.51 ± 0.19	27.08 ± 0.17
one good $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	20.7 ± 0.5	0.0130 ± 0.0012	0.120 ± 0.011	4.19 ± 0.07
one good Z + one good W + one b-jet	6.09 ± 0.31	0.0044 ± 0.0003	0.079 ± 0.002	$\sim 3 \cdot 10^{-5}$
one good Z + one $t \to Wb$	5.50 ± 0.26	$(9.91 \pm 1.4) \cdot 10^{-4}$	$(1.02 \pm 0.32) \cdot 10^{-3}$	0
$100 < M(qZ) < 250$	5.32 ± 0.24	$(9.91 \pm 1.4) \cdot 10^{-4}$	$(0.76 \pm 0.24) \cdot 10^{-3}$	0

Selection	$Z(\to ll) + j_{85-150}$	$Z(\to ll) + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
HL Trigger	36.01 ± 0.15	43.20 ± 0.17	16.18 ± 0.19	24.81 ± 0.16
one good $Z^0 \to ll$, after $\Delta M(ll)$	15.25 ± 0.09	18.35 ± 0.11	6.81 ± 0.12	0.797 ± 0.29
one good $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	0.013 ± 0.003	0.27 ± 0.04	1.86 ± 0.059	0.005 ± 0.002
one good Z + one good W + one b-jet	0	0	0	0
one good Z + one $t \to Wb$	0	0	0	0
$100 < M(qZ) < 250$	0	0	0	0

3.4 The FCNC $t \to \gamma q$ decay channel analysis

In attempting to identify the FCNC $t \to \gamma q$ decay channel, similar issues relating to jet-flavour identification must be addressed. In addition, a high-quality single photon selection strategy has to be adopted in order to avoid contamination from electrons (which may produce electromagnetic clusters close to photon candidates) or secondary photons within hadronic jets. In the first part of the section, the selection of electrons, muons and photons is studied in detail, then the issue if a trigger on a single photon is better than one on a single lepton is raised. The selection of higher level objects proceeds with the reconstruction of a W boson in the transverse plane, the optimization of the b and light-jet selections and finally the reconstruction of the two top quark decays. A large set of background sources is considered and the set of cuts demonstrates efficient in rejecting most of them.

3.4.1 Trigger and Lepton/Photon pre-selection

The signal presented in this part exhibits the presence of one hard photon and one high energy W boson, which is quite easy to find from its leptonic decay. Therefore, the choice of the trigger menu is less straightforward and both a single-lepton and single-photon streams are possible. In the following, performance and optimization of the pre-selection cuts are discussed for both the two possible choices. At the end of this part, electron and photon pre-selection will be compared and the regions where one of the choices could be better are identified.

Single-lepton and single-photon triggers

If the objects delivered by the single-lepton trigger are exploited, the lepton pre-selection strictly resembles the one presented for the previous channel. The choice of the selection cuts is optimized with the same algorithm, aiming to maximize the purity and the background rejection while maintaining a good efficiency.

In the trigger simulation adopted here, at the L1 the threshold for single muons is $3\,\mathrm{GeV}/c$, while for single electron or photon (that share the feature of a large E_T deposit in the ECAL and are not distinguished at this level) it is $p_T = 23\,\mathrm{GeV}/c$. For the HLT level, the threshold is $p_T = 19\,\mathrm{GeV}/c$ for single muons, $p_T = 29\,\mathrm{GeV}/c$ for single electrons and raised up to $p_T = 80\,\mathrm{GeV}/c$ for single photon.

Selection efficiencies for the single photon stream have been measured for the signal and all the considered background, and are indicated in Tab. 3.22. The efficiencies for single-lepton stream are in Tab. 3.21.

Table 3.21: The efficiencies for the L1 and HLT single-lepton triggers, for the complete set of considered background. All values are in percent and errors are statistical only.

	Signal $t \to q\gamma$	$t\bar{t} \to l + X$	single-t	$Zb\bar{b}$	$W + j_{85-150}$
L1 Trigger	95.94 ± 1.3	48.8 ± 0.4	74.3 ± 0.2	89.4 ± 0.3	41.39 ± 0.15
HLT Trigger	68.72 ± 1.2	39.1 ± 0.4	69.8 ± 0.2	79.2 ± 0.3	33.0 ± 0.14
	$ZZ \to 4l$	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
L1 Trigger	75.0 ± 0.3	58.26 ± 0.19	67.01 ± 0.2	52.3 ± 0.3	72.0 ± 0.9
HLT Trigger	74.9 ± 0.3	58.0 ± 0.18	63.7 ± 0.2	38.2 ± 0.3	47.8 ± 0.3

Table 3.22: The efficiencies for the L1 and HLT single-photon triggers, for the complete set of considered background. All values are in percent and errors are statistical only.

	Signal $t \to q\gamma$	$t\bar{t} \to l + X$	single-t	$Zb\bar{b}$	$W + j_{85-150}$
L1 Trigger	95.94 ± 1.3	48.8 ± 0.4	74.3 ± 0.2	89.4 ± 0.3	41.39 ± 0.15
HLT Trigger	95.74 ± 1.2	30.0 ± 0.4	2.22 ± 0.2	80.1 ± 0.3	37.03 ± 0.14
	$ZZ \to 4l$	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
L1 Trigger	75.0 ± 0.3	58.26 ± 0.19	67.01 ± 0.2	52.3 ± 0.3	72.0 ± 0.9
HLT Trigger	70.9 ± 0.3	57.76 ± 0.18	64.0 ± 0.2	47.9 ± 0.3	67.3 ± 0.3

Electron pre-selection in $t \to \gamma q$ signal

Compared with the $t\bar{t} \to qZ + W\bar{b}$ signal, the $t\bar{t} \to bW + q\gamma$ channel has a factor 3 less of muons and electrons. On the other hand, the number of b quarks decaying into leptons is nominally the same, so a worse initial selection purity is expected.

Figure 3.43(left) compares the transverse momentum distribution of the triggered electrons with the 'true' information from simulated samples, along with electrons coming directly from b quark. As expected, contamination is quite large at low energies, falling to about 20% when $p_T > 15\,\mathrm{GeV}/c$. The difference between the sum of electrons from W and from b and reconstructed objects is due to other electron sources, and is equally reduced of a factor around 3. This means that the extra-electrons are likely to originate from the electrons (*e.g.*, fake tracks and bremsstrahlung followed by a photo-conversion) rather than hadrons and heavy flavour, which number is unchanged in the two channels.

The features of the signal are compared only with the background sources that will prove to be relevant in the next step: here they are the $t\bar{t} \to l + X$ inclusive production and the single-top t-channel. The discrimination between genuine electrons and jets is rather less efficient here, since the hadronic final state of the signal (one b-jet and one light-jet) is much more prone to be mimicked by these background processes.

Figure 3.44 (left) shows that a difference in the distribution of $Isol(e)$ is visible only

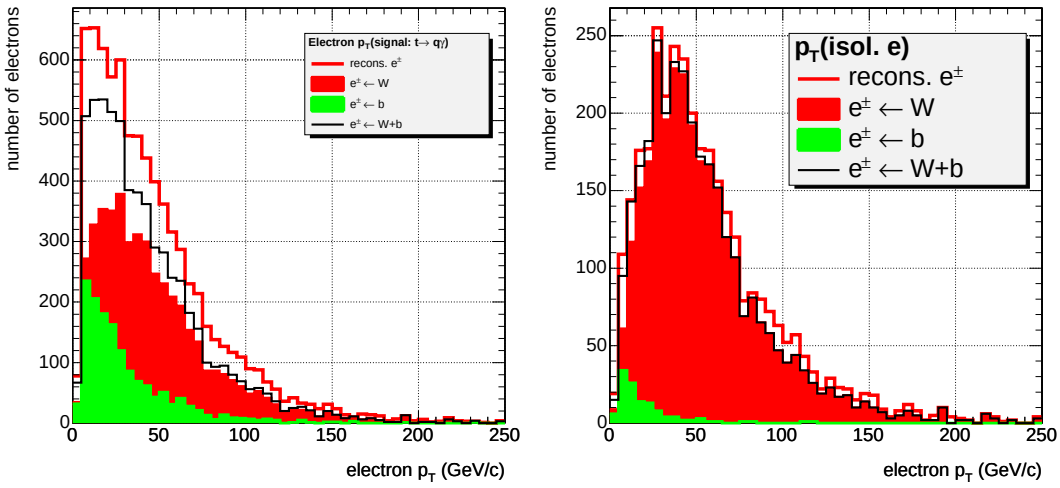

Figure 3.43: The p_T of reconstructed electrons in the $q\gamma + bW$ signal, compared with distribution of the true electrons. Comparison is shown with true electrons from W (red filled area), b-jet decay (green filled area) and the sum of two (thin solid). The gap between the sum and the reconstructed electrons has to be filled with the contribution from jets, c-decay and (mostly) photo-conversions. No isolation criteria are applied in the left plot, while the right plot considers pre-selected electrons.

for quite high values; numerical evaluation of the $N_S/\sqrt{N_B}$ demonstrates that there is no optimal isolation threshold. The choice here is to adopt the same cut of the qZ pre-selection ($Isol(e) < 0.07$), that allows to reject a part of the non-isolated electrons hidden in the multi-jet configuration of the $t\bar{t}$ background. The choice will be confirmed below by comparing the signal efficiency against the background rejection.

The strong similarity among the signal and these backgrounds exhibits even more clearly in the E_{rec}/p_{in}. The behaviour is fairly similar to that in Fig. 3.13(right), where all distributions peak around 1 with a mean and a width that are equal within few percent. Therefore, the pre-selection cut $E/p > 0.8$ optimized for the previous analysis is not altered here.

The E_{HCAL}/E_{ECAL} is a bit more sensitive to the background hadronic activity faking the electron from W boson, featuring a richer tail in the distribution (Fig. 3.44, right). Though no clear indications emerge from the optimization, the $E_{HCAL}/E_{ECAL} < 0.03$ cut may be help in rejecting a small part of $t\bar{t}$ inclusive background.

Finally, $|\Delta\eta_{in}| = |\eta(SC) - \eta(\text{extrap. track})|$ stops to be an useful quantity here, because the spread of the η difference is larger even for the signal here. In fact, an energy deposit in the ECAL may be produced even from the hard photon (preferentially at some distance from the electron track), washing away every differences with distributions in the background.

Therefore, the same quality cuts of the previous analysis are chosen as a starting working point. The most important ones, the momentum cut and the isolation level, will be indicated the former by a new efficiency/rejection curve, the latter by a comparison with a different trigger stream.

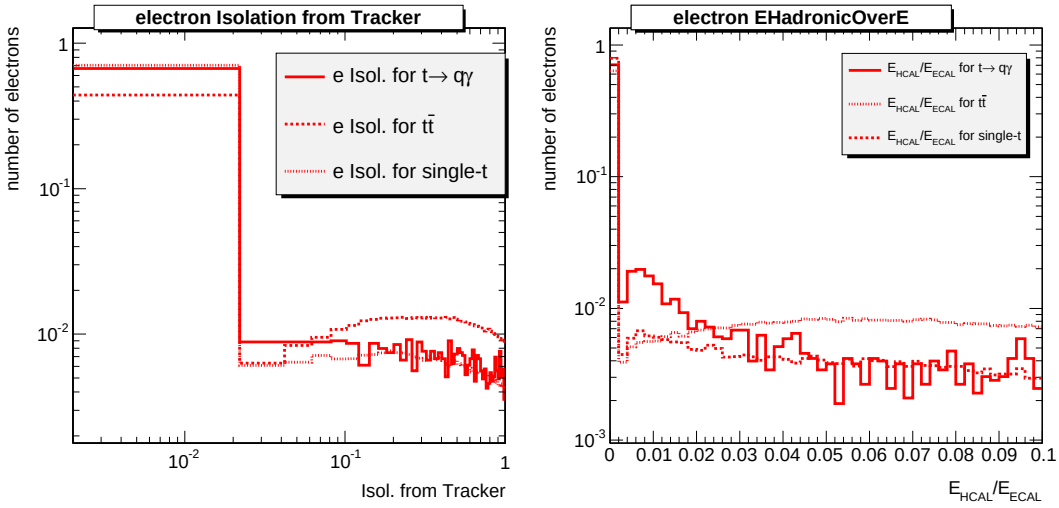

Figure 3.44: Left: Isolation in the tracker for electron from signal, single-top (dotted) and $t\bar{t} \to l + X$ (dashed) production. Right: The ratio E_{HCAL}/E_{ECAL} between energy in HCAL and ECAL for the candidate electrons. Objects with large value are likely to come from heavy flavour decay decays and have to be rejected. A normalization to the same number of events is applied on both the distributions.

Performances of the electron pre-selection and momentum cut

Figure 3.45 shows the ratio between the objects matched with the true electrons and the reconstructed one, as a function of both p_T (left) and η (right). If only electrons with $p_T > 20\,\mathrm{GeV}/c$ are considered, the behaviour of purity is quite similar to that obtained in Fig. 3.15, though limited to a 85%-90%.

The presence of a hard photon is an additional issue and may worsen the electron purity at low energy (where electrons may be well produced from a photon in the very first pixel layers) and high energy, when the electron emits a significant bremsstrahlung radiation, whose energy might be deposited in the same ECAL supercluster.

No purity dependences from η are found, just a decrease of the event rate moving from barrel to the endcap is evident, thus an increased statistical uncertainty. Contamination from b semileptonic decays, while greater than 20% without quality cuts, can be mantained below 3% only with a p_T threshold about $30\,\mathrm{GeV}/c$– that motivates the harder cut adopted already at the trigger level. The energy spectrum in Fig. 3.43(right) is the final result of the off-line pre-selection of electrons.

Even if pre-selection is a bit harder, the identification efficiency is rather good and compelling with the one evaluated in the previous analysis. In Fig. 3.46 the efficiency for a single-electron reconstruction (after the single-electron and single-muon trigger stream) is displayed in function of transverse momentum (left) and pseudorapidity (right). When $p_T(e) > 30\,\mathrm{GeV}/c$, the efficiency rises from a 35% to a 65-70%. The efficiency for electrons from b-decay ("mistag for $e \leftarrow b$"), that are the main objects to be rejected, falls below a 2% level after the p_T threshold, indicating that the isolation does a nice job.

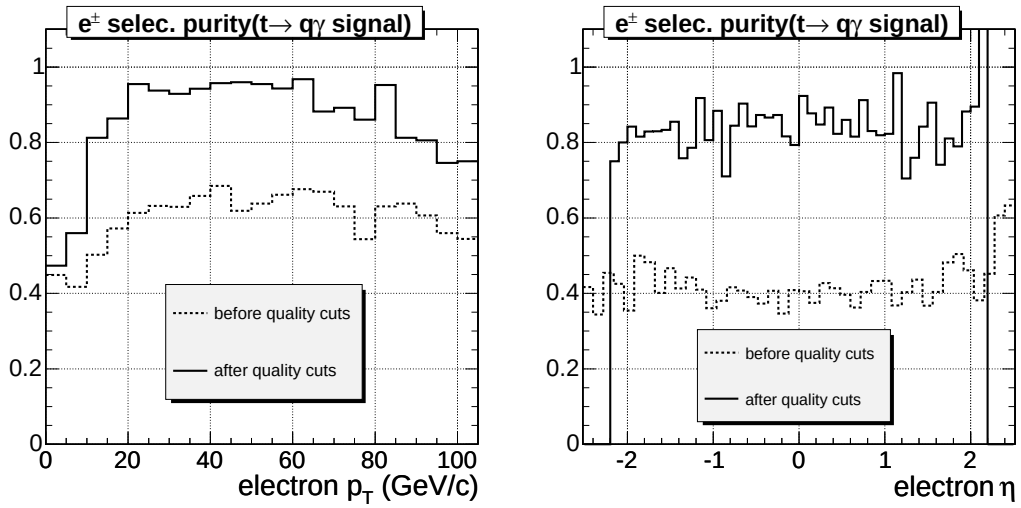

Figure 3.45: The purity of the pre-selected electrons, before and after the quality cuts, as a function of p_T (left) and η (right). The 'horn' at $\eta = +2.5$ is due to a lack of events and big error bars in this region.

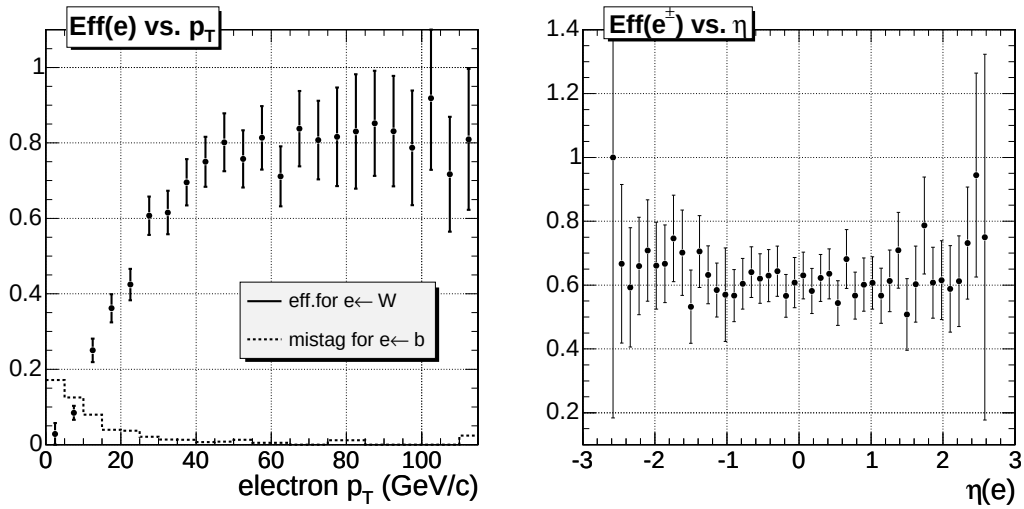

Figure 3.46: The efficiency of the electron selection after quality cuts for the $t\bar{t} \to \gamma q + W\bar{b}$ signal, as a function of electron p_T (left) and η (right). The error bars are the quadratic sum of the counting error in the histograms employed for the distribution, in each bin. The efficiency for 'spurious' electrons coming from b ("mistag for $e \leftarrow b$") is also evaluated.

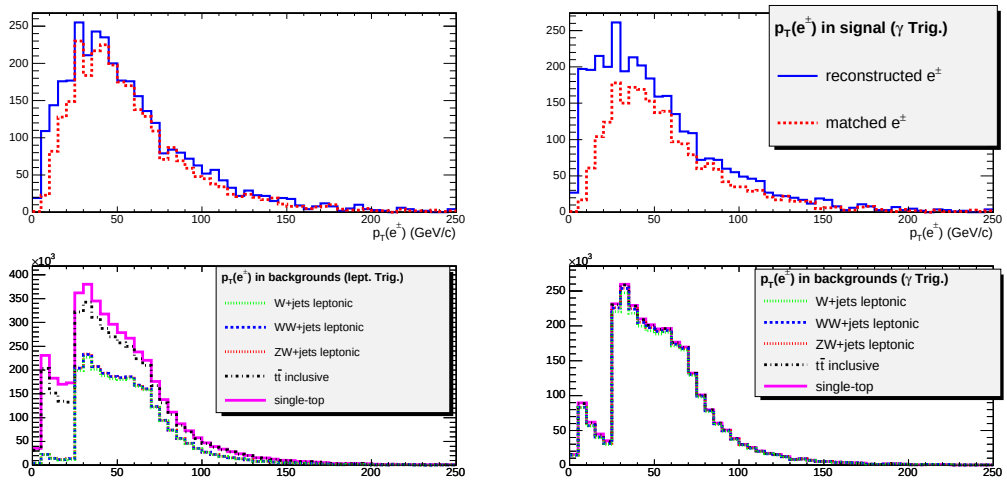

Figure 3.47: The transverse momentum distributions of $e^{\pm}$ candidates after single-electron and single-photon trigger requirements, in both the signal sample and the different background samples. The distributions for the reconstructed objects matched with the generated muons are also shown and the background samples, that are shown stacked in the lower plot, have been normalized to reflect an integrated luminosity of L=10 fb^{-1}. Account for the single-electron (left) and single-photon (right) trigger is taken.

When all the background processes are included, the situation showed in the plot in Fig. 3.47 rises. All the sources retain a sizable contribution in the sample, that is not diminished till very high energy. Only the cuts on the hard photon and the hadron part of the event (b- and light-jet) will be helpful in rejecting the di-boson and boson+jets production.

Stacked distributions for the background are significantly different between the two trigger streams. When the event is triggered by the single-photon, some of the off-line quality cuts for the electrons have to been applied yet, so a larger portion of the multi-jet background can enter. As at LO the transverse momenta of Z/W and the recoiling jets are equal and opposite, the cut on the minimum $\hat{p}_T$ implies a lower $p_T(Z)$ bound. The vector addition of this peaked distribution with the transverse momentum generated in the decay leads to the structure exhibited by the figures.

As demonstrated above, finding a lepton variable that is discriminating enough against the single and top pair production is not so easy. The leptons that constitute the major background share exactly the same origin with the signal (leptonic decay of a W that comes from a top), thus the efficiency for the signal monotonically follows that for the background. The conclusion is that the optimized selection cuts are $p_T > 30\,\mathrm{GeV}/c$ – that is the most conservative for background rejection in the preferred region, and fully retains the good efficiency for the single-electron trigger – and $Isol(e) < 0.07$. It corresponds to a signal efficiency of 0.219 ± 0.009 and a rejection of 1.135 ± 0.015.

In summary, pre-selection cuts following the single-electron triggered clone those for the $t\bar{t} \to Zq + W\bar{b}$ signal, with the only difference of the on-line and off-line momentum threshold.

Table 3.23: A summary of electron pre-selection cuts adopted in the analysis. The optimization here leads to results that are similar to the $t \to qZ$ analysis and is demonstrated to produce high purity and good efficiency.

L1	single-lepton Trigger, 23 GeV/c
HLT	single-lepton Trigger, 29 GeV/c
Isolation	$\Delta R < 0.1$, $Isol < 0.07$
Track/Cluster energy	$E/p > 0.8$, $E/p < 100$
HCAL/ECAL	$E_{HCAL}/E_{ECAL} < 0.03$
Transv. momentum	$p_T > 30$ GeV/c

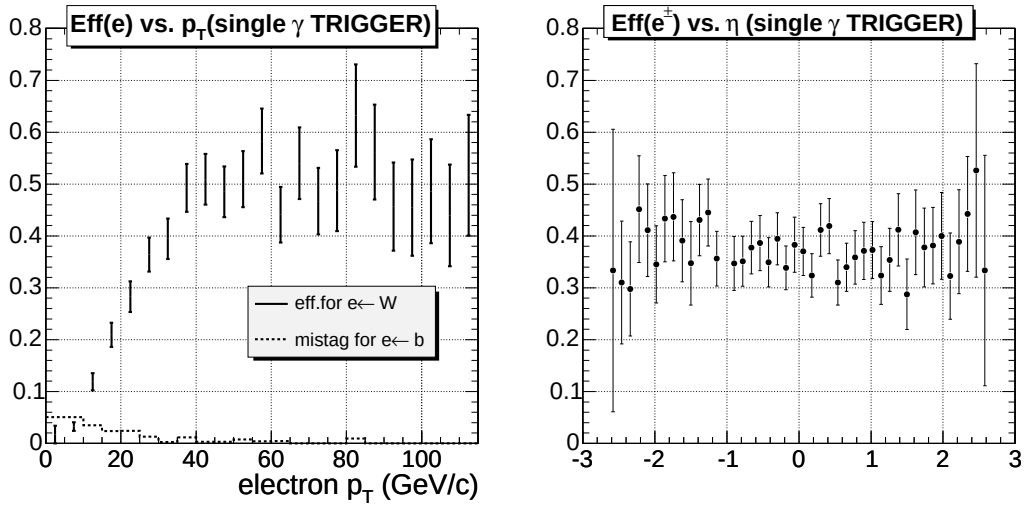

Figure 3.48: The efficiency performance of the electron pre-selection following a single-photon trigger, that can be compared with plots in Fig. 3.46. At the price of a small efficiency decrease due to the photon request, rejection against spurious electron is quite better.

They are indicated in Tab. 3.23.

When the single photon stream is adopted to trigger the event, a somewhat differences behaviour is exhibited by several variables.

While the request for an additional photon reduces the electron rate by about 75%, the purity is already good before the isolation cut, as contaminations different from the b quark decays are successfully rejected. When only electrons with $Isol(e) < 0.07$ are retained, the only spurious contribution comes from bottom decays and is below 0.8% after 30 GeV/c, therefore the purity is about 99% pratically in the whole p_T spectrum and along all η values.

The efficiency obtained with this different trigger follows the same behaviour as for the single-electron (Fig. 3.48). It reaches the saturation around 50-60 GeV/c but the value is lower, as expected from the inclusion of a well-isolated single photon. At the end of this part,

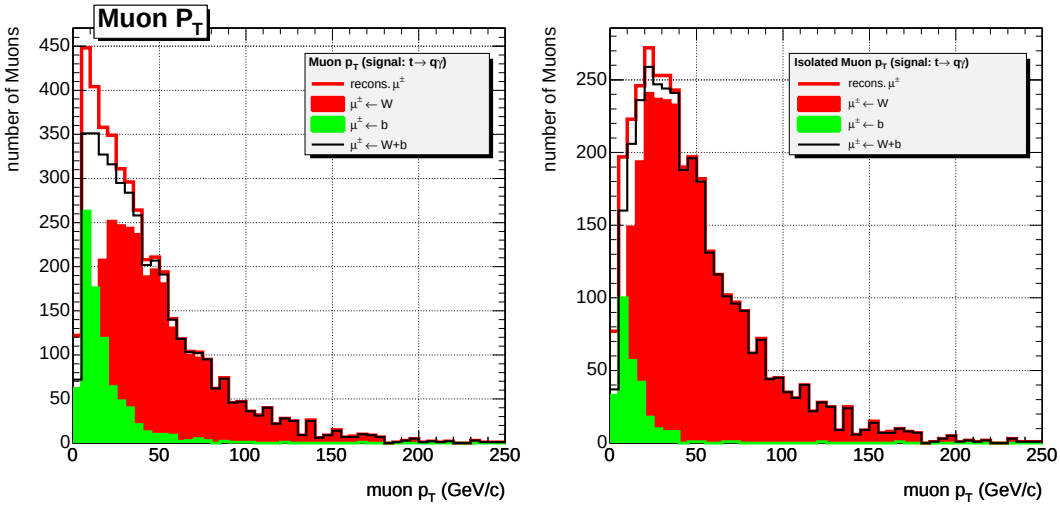

Figure 3.49: The p_T of candidate muons in the $q\gamma + bW$ signal, as they come from the single lepton trigger (left) and after the application of isolation criteria (right), compared with distribution of the true muons. Comparison is shown with true muons from W (red or dark grey), b-jet decay (green or light grey) and the sum of two (thin solid). The gap between the sum and the reconstructed objects has to be filled with the contribution from jets and c-decay and photo-conversions.

a comparison will be deployed between the $e+\gamma$ selections following the two different streams. The possibility to partially relax the isolation constrain up to 10% will be also considered.

Muon pre-selection in $t \to \gamma q$ signal

The quality of the selection for muons triggered by the single-muon stream is expressed by Fig. 3.49(left). The amount of reconstructed muons is compared with those matched with the generated ones, coming from the W in top decay or from b-jets. The rate of muons delivered from b-jets is comparable with that from the $t \to qZ$ signal. Since this signal has one lepton instead of three, the relative purity is averagely greater by a factor 3, with a 11.1% fraction for p_T <15 GeV/c .

The other contributions entering the muon selection fill the gap between the reconstructed objects and the sum of W and b quarks decay, and are significant till some tens of GeV.

The distribution of other variables involved in the muon pre-selection reflects the close similarity between the signal and the background with top. As one can guess from Fig. 3.50, nor the muon isolation in a $\Delta R < 0.1$ cone (left) neither the transverse parameter d_{xy} (right) show sizable differences in the three signals, so an optimization is not possible. Therefore, the $Isol$ parameter is set to 0.07 as in the $t \to Zq$ case, while no cuts are required on the transverse parameter. This selection has not major effect in the signals efficiency and successfully reduce the b quark contamination below the percent level, as displayed in Fig. 3.49(right). The resulting effects on purity and efficiencies of these selections are estimated below.

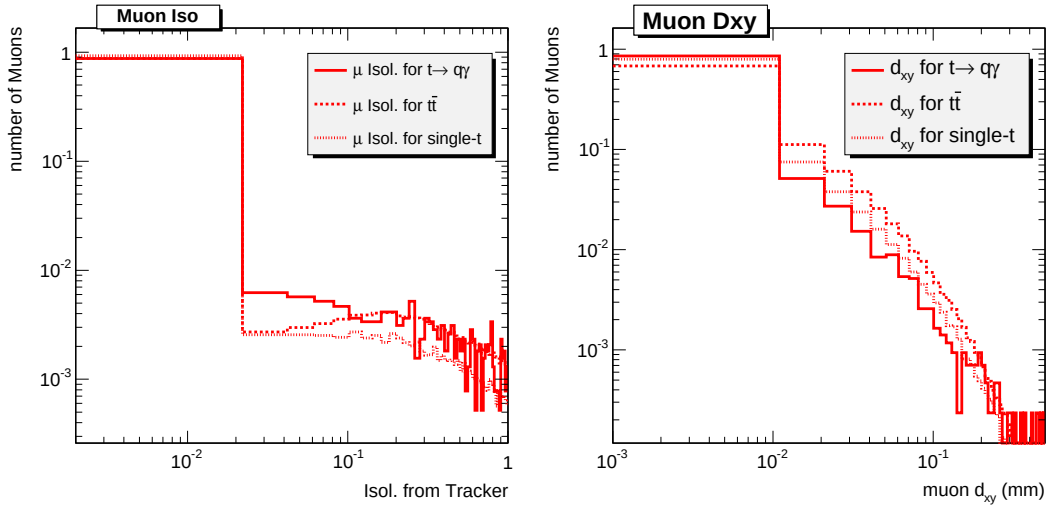

Figure 3.50: Isolation in the tracker (left) and transverse impact parameter (right) for muon from signal, single-top (dotted) and $t\bar{t} \to l + X$ (dashed) production. A normalization to the same number of events is applied.

Performances of the muon pre-selection and momentum cut

The fraction of the 'matched' muon candidates with respect to all the reconstructed objects is represented in Fig. 3.51, as a function of the transverse momentum (left) and pseudorapidity (right). Even before isolation has been applied, in the soft part of the spectrum the performance are worse, since genuine leptons are few and, below $10 \div 15\,\mathrm{GeV}/c$, the rate is comparable with that from b-decay. When their energy exceeds a $60\,\mathrm{GeV}/c$, they appear correctly selected in about 98% of the cases. The $|\eta|$ distribution reflects the purity average value, that turns out to be lower (around 90% if isolation is required) than in the $t \to Zq$ case.

On the other hand, the selection efficiency is insensitive to the different muon statistic. Performance are represented in Fig. 3.52, where also the contamination is shown for comparison.

The angular resolution is consistent with the previous analysis (if the different errors induced from the different statistic are took into consideration) and not represented here. Estimating the r.m.s. in the distribution of the different of generated and reconstructed muons, a $\sigma(\phi) \simeq 0.85\,\mathrm{mrad}$ $\sigma(\eta) \simeq 0.00087\,\mathrm{mrad}$ are found.

The p_T distributions for muon candidates after the application of both trigger and offline quality cuts, in both the signal events and in the different background samples, are represented in Fig. 3.53 for both the trigger streams. The behaviours of muon p_T for the signal and background with a top are strictly similar. When the muon efficiency is measured for different momentum thresholds and different isolation requirements, a new plot with the signal efficiency versus rejection against $t\bar{t}$ and single top can be issued. The best working results in a p_T lower threshold of $20\,\mathrm{GeV}/c$.

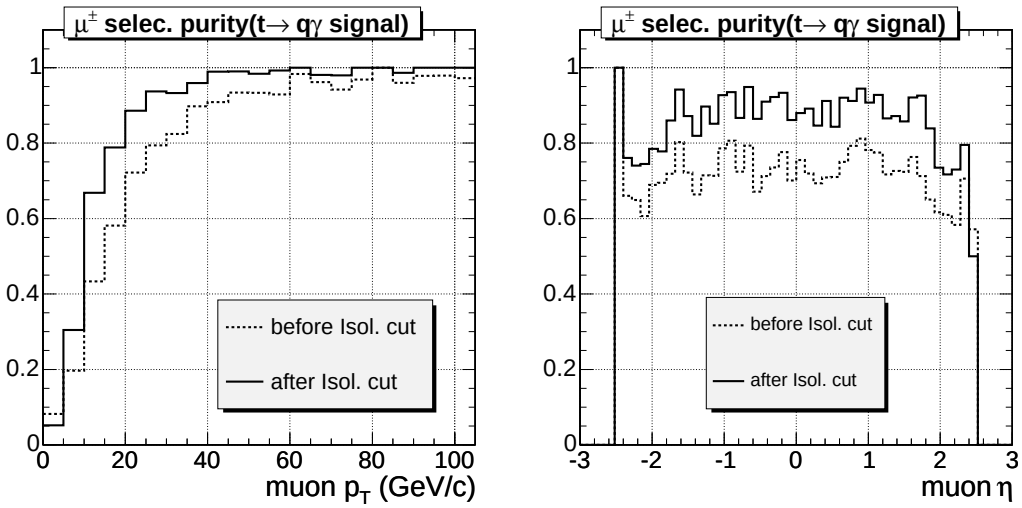

Figure 3.51: The purity of the pre-selected muons, before and after the isolation cut, as a function of p_T (left) and η (right). The error bars are the quadratic sum of the counting error in the histograms employed for the distribution, in each bin.

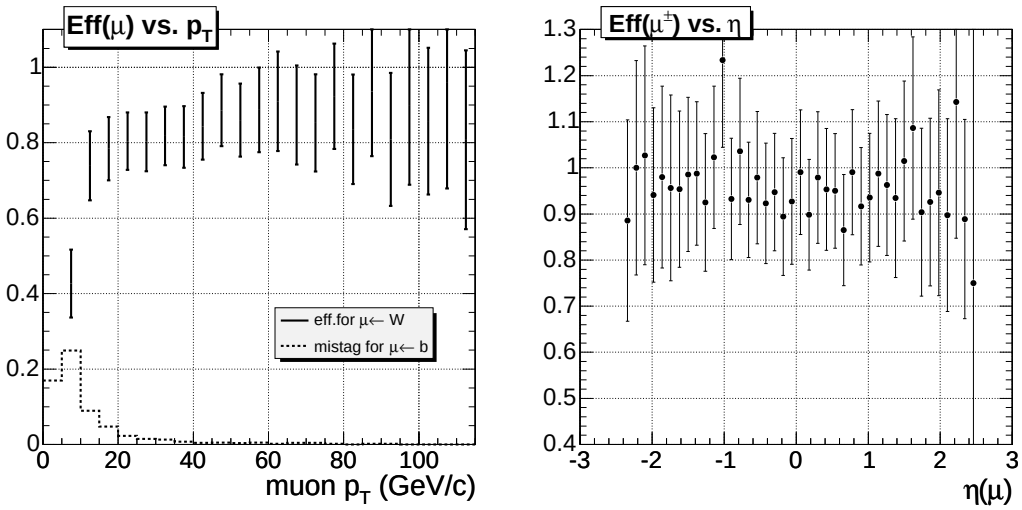

Figure 3.52: The efficiency of the muon selection after isolation cut for the $t\bar{t} \to Wb + \gamma q$ signal, as a function of muon p_T (left) and η (right). The efficiency for 'spurious' muons coming from b ("mistag for $\mu \leftarrow b$") is also evaluated.

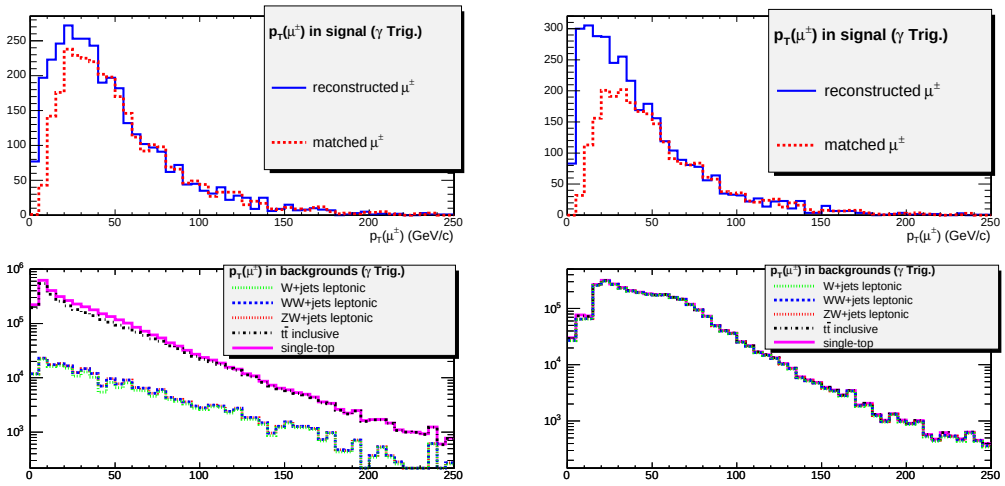

Figure 3.53: The transverse momentum distributions of $\mu^\pm$ candidates after single-muon (left) and single-photon (right) trigger requirements, in both the signal sample and the different background samples. The distributions for the reconstructed objects matched with the generated muons are also shown and the background samples, that are stacked in the lower plot, have been normalized to reflect an integrated luminosity of L=10fb^{-1}.

Table 3.24: A summary of muon pre-selection cuts adopted in the analysis. Each choice comes from the search for an optimization and is demonstrated to produce high purity and good efficiency.

L1	Single-lepton Trigger, 14 GeV/c
HLT	Single-lepton Trigger, 19 GeV/c
Isolation	$\Delta R < 0.3$, $Isol < 0.07$
Transv. momentum	$p_T > 20$ GeV/c

The Table 3.24 summarizes all the pre-selection cuts for the electron in the $t\bar{t} \to Zq + W\bar{b}$ signal.

Pre-selection of photon candidates

The baseline choice for the reconstruction of the high E_T photon coming from the FCNC top decay is to rely on a single-lepton stream, then apply the off-line selection algorithm sketched in the previous chapter (Sec. 2.3.1). Further developments to such a selection can be devised, aiming to maximize the selection purity and rejection power.

Figure 3.54(left) is the transverse energy distribution of all the photons reconstructed in this way compared with the ones matched with photons coming directly from the top quark. If photon clusters in the electromagnetic calorimeter are too close to the electron and muon candidates tracks ($\Delta R > 0.3$), they are likely to be originated from bremsstrahlung photons by charged particles, thus discarded.

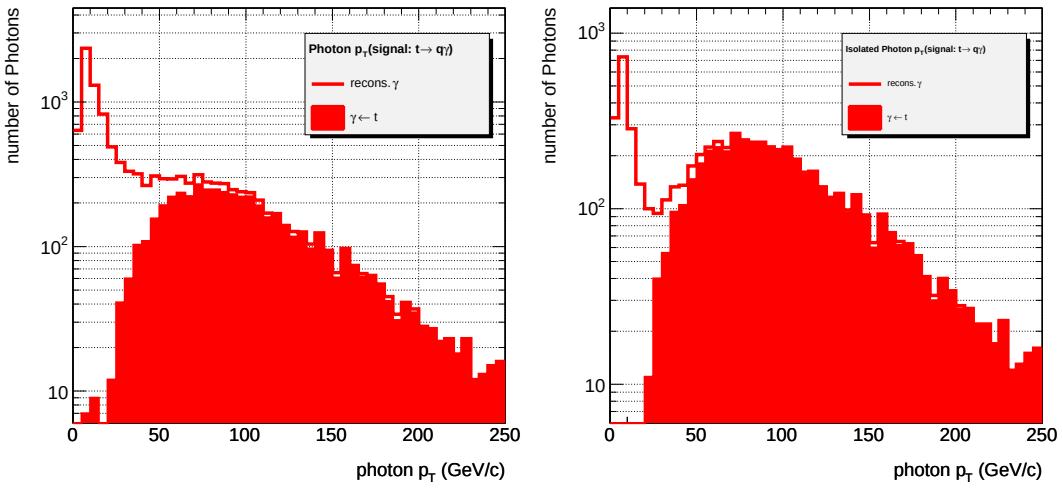

Figure 3.54: The p_T of reconstructed photons in the $q\gamma + bW$ signal (open area), compared with distribution of the interesting photon coming from top (filled area). No isolation criteria are applied in the left plot, while on the right photons are isolated.

Reconstructed objects with an energy of few p_T are mostly something different from prompt photons: among them, photons from π^0 and η^0, bremsstrahlung emissions, neutral hadrons. Therefore, γ's from top anomalous decay are nothing else than the high energy tail ($E_T > 70 \div 80$ GeV) in the distribution of the total amount of photons. If only objects with $E_T > 40$ GeV are selected, about 18% of contamination is still there.

When these high transverse energy photons are considered, the main polluting contribution is again due to jets, that may produce neutral hadrons or excited particles with radiative decays. While these photons from secondary processes are normally embedded in a dense track bundle, prompt photons from top decay are well isolated. Once again, isolation may be a powerful variable to increase both the purity of the selection and the sensitivity to background processes faking the signal.

Isolation variable $Isol(\gamma)$ for photons is here defined as the sum of the p_T of all good tracks (p_T >0.9 GeV/c and number of hits > 4) which lie within a cone of radius $\Delta R = 0.3$ around the electromagnetic cluster. If the sum is less than a given fraction of the transverse energy of the photon candidate, then it is considered to be isolated.

As for all the reconstructed objects, the optimization for a background rejection is scrutinized only for the most relevant sources. The optimization procedure is performed again by looking for a maximization of the $N_S/\sqrt{N_B}$ ratio in the photon isolation (for the signal and single-top and top pair inclusive production background). The $Isol(\gamma)$ behaviour indicates $Isol(\gamma) = 0.010 \pm 0.002$ as a best value for this threshold. When this cut is assumed, the result in Fig. 3.54(right) appears and shows how the reconstruction purity improves. The percentages of mis-matched photons are less then 4.5% for E_T >40 GeV and less then 2% for E_T >90 GeV . These results are presented in a different form in Fig. 3.55: here the selection

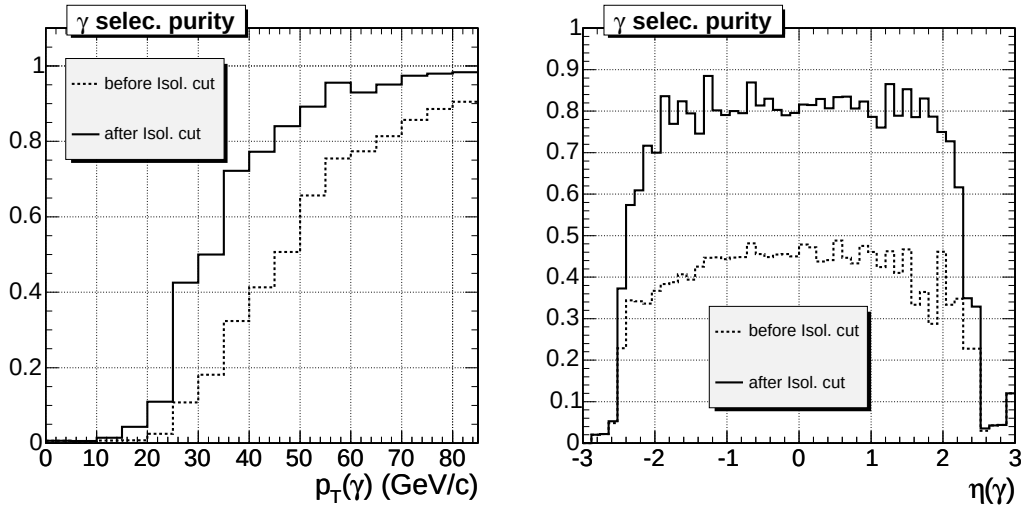

Figure 3.55: The selection purity for the γ selection, after the single-electron trigger. The photon is reconstructed as detailed in the previous chapter and the off-line cut $Isol(\gamma)$ <0.01 is here applied. The sensible improvement induced by the isolation is showed as a function of $p_T(\gamma)$ (left) and $\eta(\gamma)$ (right).

purity, defined as the fraction of matched objects in the reconstructed photon sample, is evaluated as a function of photon transverse energy (left) and photon pseudorapidity (right).

Such plots indicate that, while without isolation only the 75% of photon with E_T >60 GeV are matched correctly and the maximum purity is reached for very high emergies, isolation allows a nice $90 \div 95\%$ after already 60 GeV to be reached. The benefits of photon isolation are visible also from the running with η that, since variations with pseudorapidity seem not relevant, can be used as an average estimate of the purity. The drops that have been found for $|\eta| > 2$ are considered not significant, as they involve a small portion of the reconstructed photon and have large error bars. Thus a more severe isolation cut in the endcap region, though possible, seems no mandatory.

The efficiency for the hard photon reconstruction (selected from the standard single-lepton trigger stream) may be evaluated from the plot in Fig. 3.56, where the dependencies from transverse energy and pseudorapidity are explicited.

If only the photon harder than 60 GeV/c and with $|\eta(\gamma)| < 2.0$ are retained, it turns than efficiency (that is, the ratio of the selected isolated photons over all the photons delivered by the reconstruction) nicely exceeds 90%.

When the single-photon trigger is exploited, a quite different scenario presents. Even if no off-line isolation is applied to the photon delivered by the stream, the comparison with the generated objects shows that purity is very good down to 50 GeV . Imposing the same isolation cut as in the lepton trigger, one can verify that performance is only mildly improved. The major advantage of this quality cut is to clean from mis-identification in the central region. Here radiation of hard photon not originating from top quark is more probable, and in the

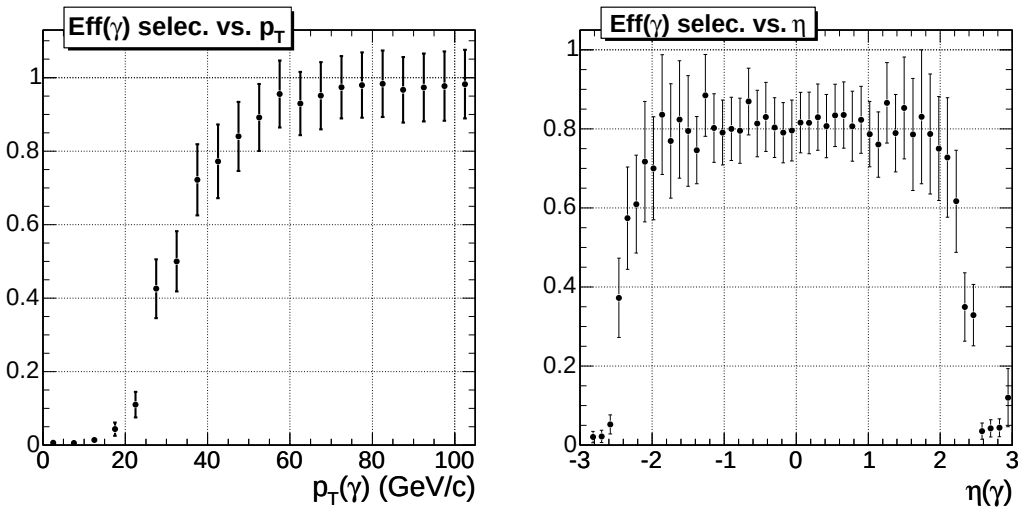

Figure 3.56: Selection efficieny for the photon after the single lepton trigger, the standard reconstruction procedure and the additional off-line isolation requirements. Performance are presented as a function of p_T (left) and η (right).

$|\eta| < 1.2$ region the purity may improved from 55% to 80%. Figure 3.57 may be compared with Fig. 3.56 to show how similar is their behaviour.

The cumulative plot showing all the background contribution is presented in Fig. 3.58. Most of the background sources may enter the sample because of a mis-identification of the photon with a lepton – mostly an electron. As expected, this occurs only for the softest particles, while after $p_T(\gamma) = 60 \div 70\,\mathrm{GeV}/c$ the effect seems successfully eliminated. Therefore, the selection for a quite hard photon demonstrates a powerful tool to reject most of the background processes. As it will shown below, this property is fully preserved (and possibly enhanced) where the single-photon trigger is exploited.

Real hard photons found in all the backgrounds drop exponentially with the energy, while the signal spectrum has a lighter slope. How to profit from this feature is showed in Fig. 3.59, where the signal and the top sources photon spectrum is plotted on a same number of events basis. From $p_T(\gamma) \simeq 50\,\mathrm{GeV}/c$, where photon identification starts to be affordable, the background continues to steadily decrease, while the signal has the bulk of its distribution. Thus a transverse energy cut seem instrumental in rejecting the radiative part of the background. Optimization procedure suggests a value between 95 and $100\,\mathrm{GeV}/c$.

In summary, the analysis for the off-line reconstructed photons suggests that the best properties – both for the reconstruction quality and background contamination – are obtained with $Isol(\gamma) < 0.01$ and $p_T(\gamma) > 95\,\mathrm{GeV}/c$.

Comparison between single-l and single-γ trigger

At this level, results are mature enough to allow a comparison between the yields provided by the leptonic and the photonic High Level Trigger. The issue one should establish is whether

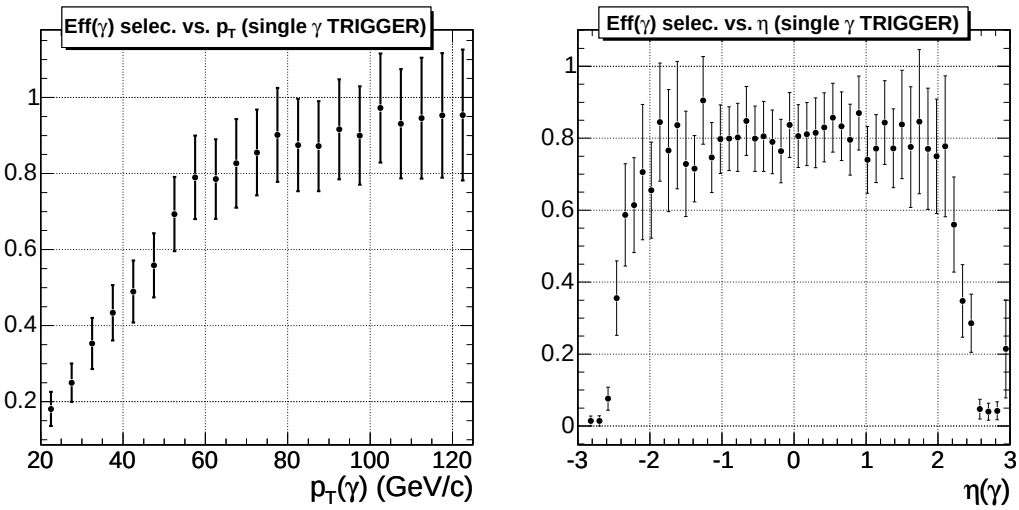

Figure 3.57: Selection efficiency of the γ in the single-photon stream, after the isolation. Performances show no relevant difference from the standard lepton stream, so the question of which of the two is preferable raises.

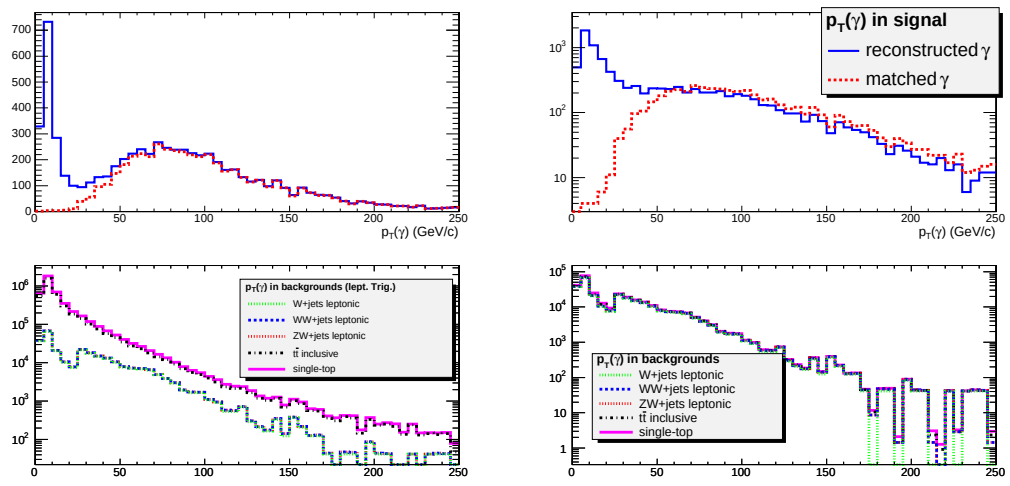

Figure 3.58: The distribution of the transverse momentum of photon candidates in both the signal sample and the different background samples, after the single-lepton (left) and single photon trigger (right).

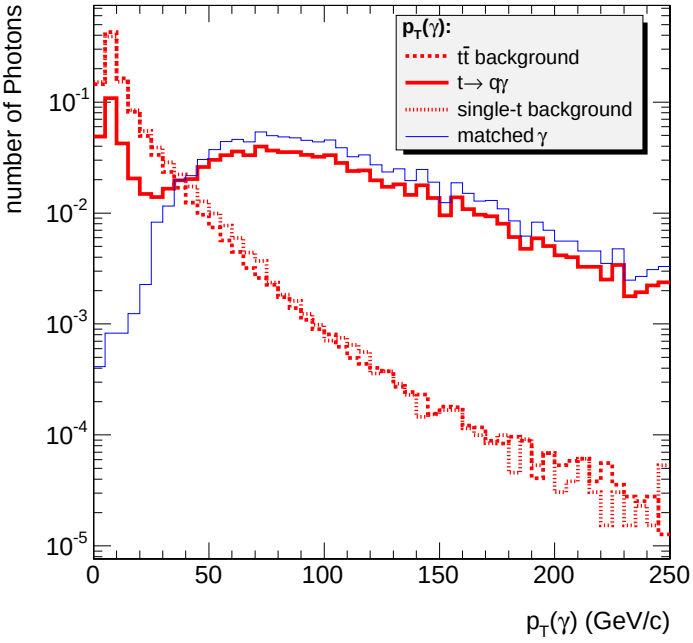

Figure 3.59: A comparison of the photon transverse energy distribution, focused on the signal and the two relevant backgrounds with a normalization to the same event number. Also the p_T for the photon matched with a generated γ (ΔR <0.3) is shown. The threshold for the optimal background suppression is found in the $95 \div 100$ GeV region.

the efficiency of the cascade cuts required up to this level is better in one case or in the other, and where the background minimization is stronger. Answering the question needs to consider both the HLT and the off-line efficiencies. If $\epsilon_{HLT}(1\text{ lept})$, $\epsilon_{HLT}(1\gamma)$ are the trigger efficiencies for two streams and $\epsilon^{l/\gamma}(1\text{ lept}|\text{lept cut})$, $\epsilon^{l/\gamma}(1\ \gamma|\gamma\text{ cut})$ are the efficiencies of the off-line pre-selections on objects delivered by the single-lepton/single-photon triggers, the yields to be compared turn out to be:

$$\epsilon^{l}_{TOT}(l+\gamma) = \epsilon_{HLT}(1\text{ lept}) \times \epsilon^{l}(1\text{ lept}|\text{lept cut}) \times \epsilon^{l}(1\gamma|\gamma\text{cut}) \tag{3.10}$$
$$\epsilon^{\gamma}_{TOT}(l+\gamma) = \epsilon_{HLT}(1\gamma) \times \epsilon^{\gamma}(1\gamma|\gamma\text{cut}) \times \epsilon^{\gamma}(1\text{ lept}|\text{lept cut}).$$

While for each stream the lepton and photon cuts have been verified to be fairly independent, it is not possible to assume that, for instance, $\epsilon^{l}(1\text{ lept}|\text{lept cut}) = \epsilon^{\gamma}(1\text{ lept}|\text{lept cut})$, since the off-line cuts operate on samples that in principle are not the same. Therefore, results for these global efficiencies are quoted below, also indicating when the lepton is a e or a μ. As both photon and lepton efficiencies have demonstrated to be quite sensitive to the photon and electron isolation, all values are evaluated for 2(3) different upper cuts of the $Isol$ variable in the photon (lepton) case. The muon has very few chances to be mis-identified with a photon or an electron, so its efficiency is rather independent from the stream and the $Isol(\mu)$ parameter has been fixed. In Tables 3.25 and 3.26 the signal(background) efficiencies

Table 3.25: The selection efficiencies for a signal $t\bar{t} \to (\gamma q)(W\bar{b})$, when only a hard lepton and a hard photon are required. Here the trigger is given by the lepton and the photon is selected off-line. Estimates are repeated for electron and muons and values are in percent. The same selection repeated for the $t\bar{t}$ background is in parenthesis.

	$Isol(e) < 0.01$	$Isol(e) < 0.1$
$Isol(\gamma) < 0.015$	$e:$ 4.916(0.118) $\mu:$ 4.917(0.118)	$e:$ 4.984(0.117) $\mu:$ 4.997(0.118)
$Isol(\gamma) < 0.03$	$e:$ 4.952(0.118) $\mu:$ 4.952(0.118)	$e:$ 5.020(0.117) $\mu:$ 5.033(0.118)
$Isol(\gamma) < 0.05$	$e:$ 5.291(0.128) $\mu:$ 5.410(0.128)	$e:$ 5.364(0.128) $\mu:$ 5.088(0.129)
$Isol(\gamma) < 0.1$	$e:$ 5.004(0.158) $\mu:$ 5.004(0.158)	$e:$ 5.073(0.167) $\mu:$ 5.114(0.168)

after the trigger, lepton pre-selection, photon pre-selection cuts are reported (in percent) for the some relevant isolation combinations.

A couple of observations may be drawn from these results:

- as expected, both the signal and background rates increase with softening the isolation, as a greater amount of event is included. When the $Isol(\gamma)$ is large, spurious contributions around the photon cluster could be collected, so the efficiency rises more slowly, or even decrease;

- in the signal, the muon and lepton rates are equal within the statistical errors. This is true in the background with leptonic trigger also. When the photon trigger is switched on and the additional photon is required, the muon rate strongly decreases. The electron rate, that here appears to be richer by a factor ~ 6, can be ascribed to the mis-identified photons.

Due to the high rates both in the signal and in the $t\bar{t}$ background at this level, statistical errors affect only the last significant digit.

In summary, best results for the selection of the electromagnetic component of the signal are offered from the single-photon trigger and from a pre-selection of electrons and photon that predicts isolation cuts $Isol(e) < 0.1$ and $Isol(\gamma) < 0.015$. Consequently, this is the trigger stream and pre-selection cuts that has been adopted here. The reconstruction of high level objects, detailed in the following, is performed with the lepton and photon delivered by this stream.

Table 3.26: The same that the previous selection efficiencies, here with a trigger from single-photon and the lepton found off-line. The comparison of signal and background efficiencies allow to compare the two streams and to properly choose the isolation cuts.

	$Isol(e) < 0.01$	$Isol(e) < 0.1$
$Isol(\gamma) < 0.015$	$e:$ 5.089(0.161)	$e:$ 5.159(0.161)
	$\mu:$ 5.089(0.030)	$\mu:$ 5.165(0.032)
$Isol(\gamma) < 0.03$	$e:$ 4.943(0.174)	$e:$ 5.011(0.173)
	$\mu:$ 5.049(0.033)	$\mu:$ 5.124(0.034)
$Isol(\gamma) < 0.05$	$e:$ 5.218(0.199)	$e:$ 5.290(0.198)
	$\mu:$ 5.219(0.038)	$\mu:$ 5.296(0.039)
$Isol(\gamma) < 0.1$	$e:$ 5.221(0.242)	$e:$ 5.293(0.241)
	$\mu:$ 5.292(0.046)	$\mu:$ 5.371(0.047)

3.4.2 The $W^{\pm}$ boson reconstruction

The W boson decaying in a lepton and a neutrino is the first final state object to be reconstructed in the analysis. Therefore, the impact of the background sources is much different from the previous case, because no other high level object is requested before it, and all the process featuring a hard lepton and enough missing energy are supposed to enter the selection.

Before combining the lepton and the transverse missing energy to form a transverse invariant mass bound, a closer look to the latter is necessary.

Constraints on the missing transverse energy

As demonstrated from the plots below, a discrimination based on the missing energy is not possible and most pernicious background cannot be disentangled via this variable. The comparison of signal and all the considered background in Fig. 3.60 (left) testifies that all the sources share a similar behaviour and are just a tiny excess on top of the huge W+jets. In Fig. 3.60 (right) this situation is clarified by showing the three backgrounds that have some chances to survive till the last cut.

Missing energy spectrum of $t\bar{t}$ inclusive events tends to widen back to the low energy region, since they contain a much larger hadron activity than the previous analysis, so a larger production of not very hard neutrinos. This effect causes the approaching of E_T^{miss} distribution in $t\bar{t}$ to the one for single-top (where only one W is produced), in fact making not possible a clear separation of the two processes. No preferred value in respect of background minimization are shown: consequently, the soft cut $E_T^{miss} > 25$ GeV is issued, that is instrumental to avoid neutrinos from other spurious sources.

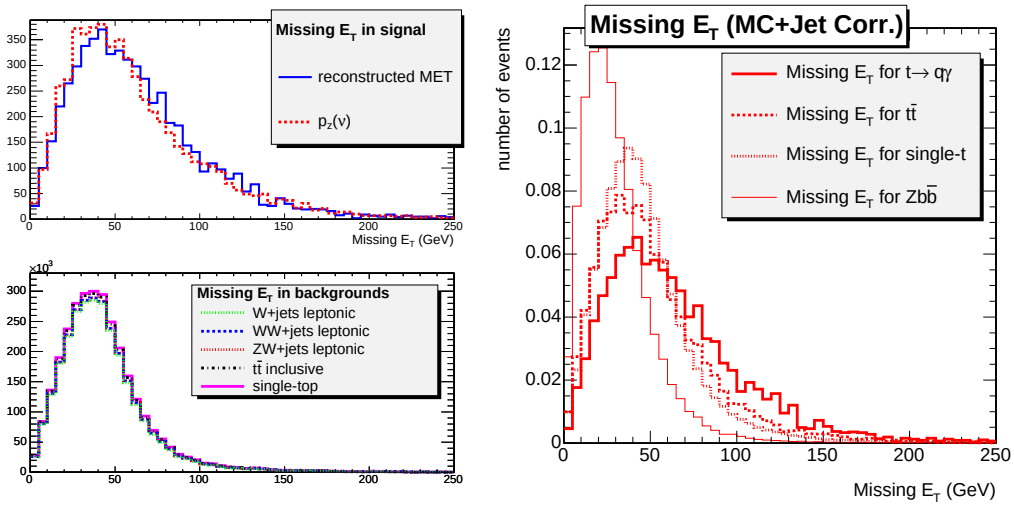

Figure 3.60: Left: The comparison of the E_T^{miss} for the signal and relevant background. In the lower plot, rates are scaled at the $10\,\mathrm{fb}^{-1}$ integrated luminosity. Right: the same distribution from signal and the two most important background sources, when normalized to the same number of events.

Constraints on the W mass

The amount of missing energy is combined with the electron and muon from the photon trigger stream and pre-selection cuts, then a transverse invariant mass bound is issued.

Distributions computed from Eq. 3.4, with the additional requirement to keep the lepton-missing transverse energy combination with the invariant mass closest to the mass of the $W^{\pm}$, are plotted in Fig. 3.61 for electron (left) and muon (right). The three-fold increase of statistic with respect to the previous analysis allows to better reproduce the expected shape of the distribution and its smoothed edge is found closely to the generated mass. Again, it is compared with the transverse mass calculated with the E_T^{miss} and an electron/muon tightly close to a true electron/muon coming from the W. The match with the generated objects is found to be very good in both cases.

If an event is found to contain more than one object consistent with a $W \to l\nu$, it is skipped.

By inquiring the kinematic distributions of these reconstructed $W^{\pm}$, statistic seems reach enough to allow a deeper scrutiny. A clear different distributions in W transverse momentum are found. As shown in Fig. 3.62, where the variable $p_T(W)$ is compared by referring to the same number of events, W boson from the single-top seem significantly softer, especially in the muon case where selection efficiency is larger and mis-identifications are much less likely to occur. The effect can be ascribed to the fact that the top quark here recoils against a light quark, whose significant part of momentum is forward or backward and not transverse.

A $p_T(W)$ cut can thus be suggested, and a numerical optimization has been performed to design the signal region with the best statistical properties. Such a procedure gives the lower thresholds $p_T(e\text{-}E_T^{miss}) = 65\,\mathrm{GeV}/c$ and $p_T(\mu\text{-}E_T^{miss}) = 50\,\mathrm{GeV}/c$ to be applied to the signal.

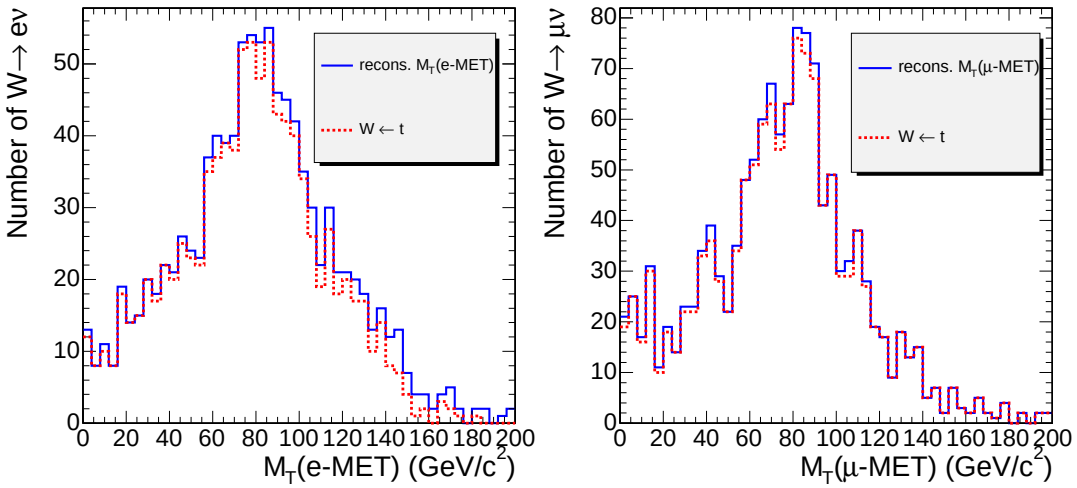

Figure 3.61: The transverse mass distribution of the W in the signal, intended as the transverse invariant mass combination of a hard lepton coming from the interaction point and large transverse energy. In both plots, the expected Jacobian shapes are nicely reproduced.

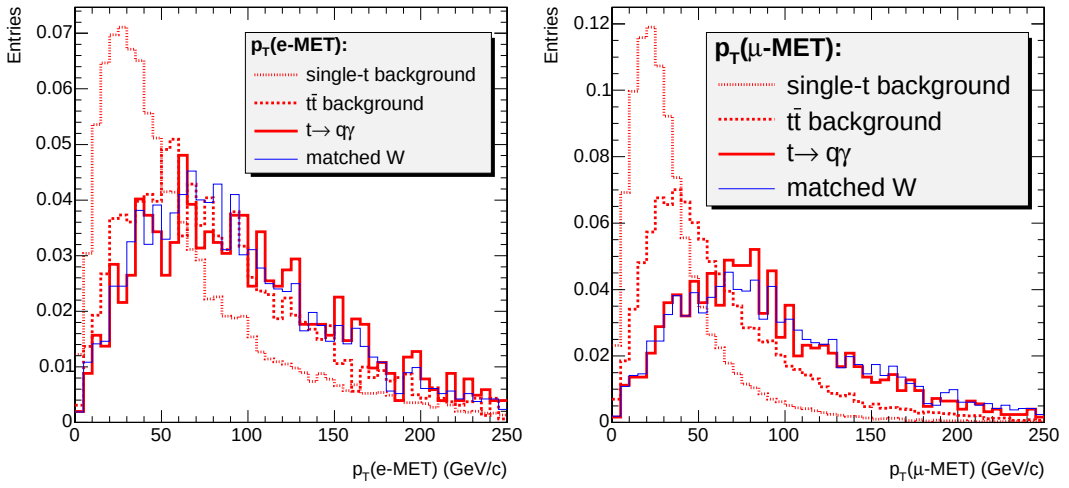

Figure 3.62: The transverse momentum of the reconstructed W bosons, evaluated on $p_T(e\text{-}E_T^{miss})$ (left) and $p_T(\mu\text{-}E_T^{miss})$ (right). Distributions from single-top production is fairly separated, so an additional p_T cut can be suggested to reduce this important background.

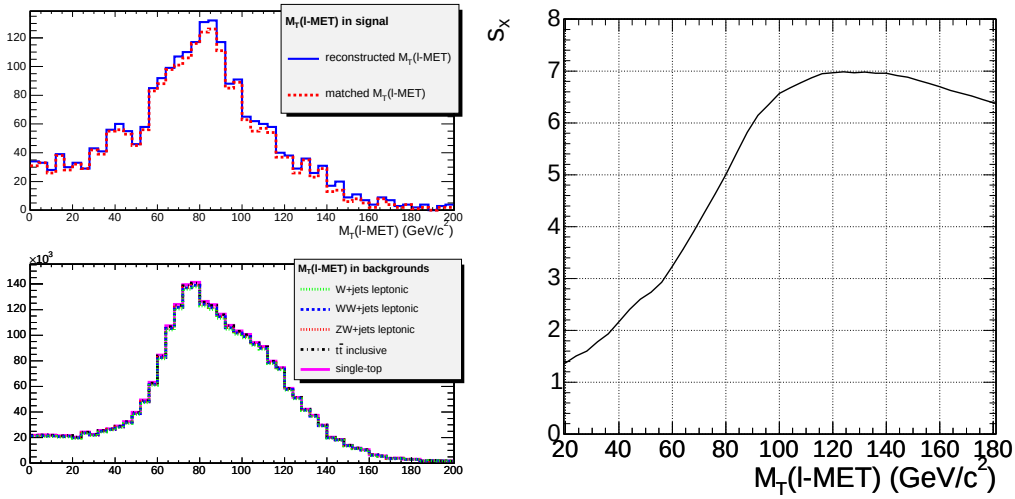

Figure 3.63: Left: The distribution of the transverse invariant mass of the $e^{\pm}$ or $\mu^{\pm}$-missing energy combination in the signal sample and in the different background samples. The background distributions have been normalized assuming an integrated luminosity of L=10 fb^{-1}. Right: The significance s_x (related to the signal efficiency and the number of surviving background) as a function of W invariant mass cut. The large smooth region around maximum testifies the robustness against missing energy fluctuations.

They result in a suppression of the single-top contribution by a factor of 0.400 and 0.244 in the $M(e\text{-}E_T^{miss})$ and $M(\mu\text{-}E_T^{miss})$ respectively that, thanks to the better muon efficiency, corresponds to a 0.291 cumulative reduction in $M(l\text{-}E_T^{miss})$. If the reduction on signal is compared with that on the $t\bar{t}$, single-top and also W+jets processes, one found that statistical power is enhanced by about 9% with respect to no momentum cuts. Moreover, other processes featuring softer leptons combined with smaller missing energy will be effectively ruled out.

It is important to observe that subsequent selections can be assumed to be independent from this kinematic constrain, since the light-jet is un-correlated to the W boson and the photon has been already selected from the trigger. The requirement of this quite 'boosted' objects could influence the energy of b-jets that will be chosen subsequently, forcing the selection to prefer jets with larger momentum. Such effect is fairly small, and pushes further in the direction to favour the signal with respect to the single-top.

If the selection is repeated for all the relevant background, plots as the one in Fig. 3.63(left) are obtained. Compared with the W+jets, other backgrounds but the inclusive $t\bar{t}$ have a very modest rate; the latter will become largely more important only when the boson plus jet processes will be completely suppressed. This will be possible only with the very last cuts.

Even in this case, significance of the selection can be slightly improved if an upper cut in the transverse mass is applied. In Fig. 3.63(right), the estimator s_x of the significance is evaluated between the signal and the $t\bar{t}$ and single-top process, shifting the upper threshold in $M_T(l\text{-}E_T^{miss})$. The better value for this threshold is found between 120 and 140 GeV/c^2. While the enhancement of significance is quite modest with respect to not apply any mass

Table 3.27: The efficiencies for the reconstruction of a W boson in the leptonic decay channels, from all the considered background. Kinematic cuts discussed in the test are applied to leptons and invariant mass. All values are in percent.

Selection	Signal $t \to q\gamma$	$t\bar{t} \to l + X$	single-t	$Zb\bar{b}$	$W + j_{85-150}$
one $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	24.2 ± 0.5	8.59 ± 0.14	0.60 ± 0.06	8.64 ± 0.07	5.99 ± 0.04
Selection	$ZZ \to 4l$	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	2.88 ± 0.03	4.72 ± 0.03	4.77 ± 0.04	4.53 ± 0.05	7.87 ± 0.05

cuts, choosing this point sets the mass selection in the point where it is more robust against systematic variation. Setting this work-point, the final sensitivity of the selection is worsened by a (relative) 10% only if the mass resolution would be known not better than 20%, that is a very pessimistic scenario.

All the resulting efficiencies are detailed in Tab. 3.27. With respect to what has represented in Fig. 3.63(left), here also the requirement on the W transverse momentum is inserted.

3.4.3 Light-jets and b-jets specific analysis

The procedures to tag b-jets coming from the standard top quark and to identify light-jet from the FCNC decay, have been detailed earlier, and applied to the analysis of the $t\bar{t} \to (qZ)(\bar{b}W)$ signal. The same selection is repeated for the $t\bar{t} \to (q\gamma)(\bar{b}W)$ signal and corresponding background, and the results are discussed.

Selection of the final state b-jets

The transverse momentum distribution of the b-tagged jets, selected by applying the algorithm explained in Sec. 3.2.3 to the event with one W boson reconstructed, strictly clones what was displayed in Fig. 3.36(left). The most relevant background comes from the abundant b-jet production by the other top quarks in the event, coming both from single and pair production. Though the $p_T(b\text{-jet})$ spectrum for single top (t-channel) tends to be slightly softer, no big differences exists – as Fig. 3.64 (left) demonstrates – so an optimization is not possible and the standard threshold of $50\,\mathrm{GeV}/c$ is not modified.

On the other hand, the multiplicity distribution (reported in Fig. 3.64 (right)) for these jets shows interesting features. Having included in the selection a so large component of the inclusive $t\bar{t}$ production, reflects in the fact that the fraction of events with two b-tagged jets are quite larger in this source that in the signal. Consequently, the veto for a second tagged jet is powerful in strongly reducing the $t\bar{t} \to (Wb)(W\bar{b})$ rate. Moreover, some distinction is possible between this process and the single-top production, since the latter behaves quite

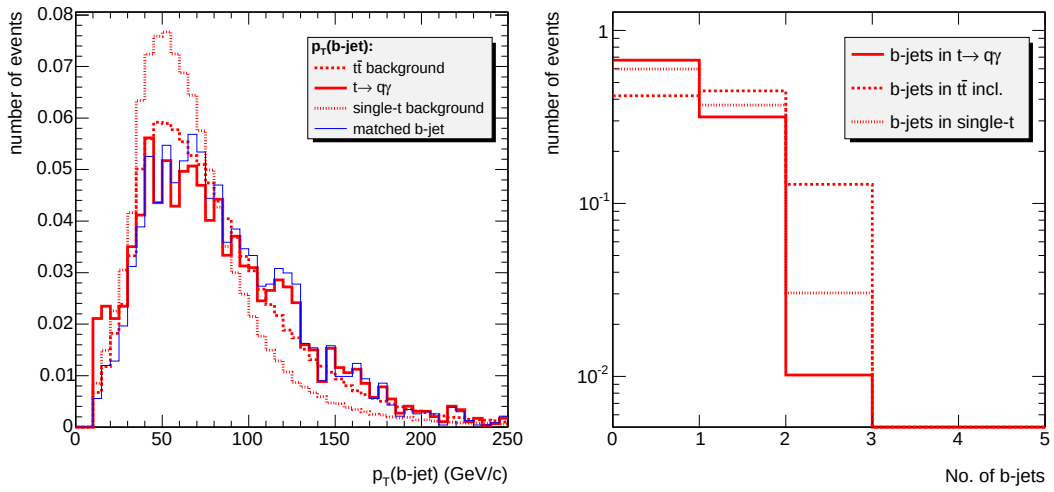

Figure 3.64: Left: A comparison of transverse momentum of selected b-jets in signal and two relevant backgrounds. Also matched jets are shown and comparison is on the same event number. Right: the relative distribution of the number of b-jets, once the presence of the W boson is asked for.

Table 3.28: Selection efficiencies when the W selection is supplemented by the request for one b-tagged jets. Not all these backgrounds have been represented in the plots. All values are in percent.

Selection	Signal $t \to q\gamma$	$t\bar{t} \to l + X$	single-t	$Zb\bar{b}$	$W + j_{85-150}$
one $W \to l\nu$ and one b-jet	6.8 ± 0.3	3.17 ± 0.09	0.222 ± 0.002	1.22 ± 0.02	0.089 ± 0.005
Selection	$ZZ \to 4l$	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one $W \to l\nu$ and one b-jet	0.024 ± 0.003	0.097 ± 0.005	0.217 ± 0.009	$(9.6 \pm 2.3) \cdot 10^{-3}$	0.183 ± 0.008

similarly to the signal and the fraction of double-tagged events are a factor ~ 15 lower than $t\bar{t}$. Although this distinction could be useful in separating the background sources each other (that is mandatory when background will be estimated in specific control region), it can suffer from systematic effects on the b-jet multiplicity, thus it is safer to not exploit it.

As expected, the single-top exhibits a discriminator value that is significantly lower than that from the top associated production (Fig. 3.65(left)) while the one from $t\bar{t}$ closely resembles the signal, and no further optimization are possible on this variable. Therefore, the discriminator selection chosen earlier for the best purity/efficiency/mistagging performance is left untouched. In the Fig. 3.65(right) the comparison is extended to the boson+jet processes that, because of the large cross section, may enter the selection thanks to even a small degree of mis-tagging. In fact, Z/W+jets background demonstrates to have a rate less than a factor 10 smaller than the signal, thus it will be fully rejected only after further cuts. Since these cuts will be successful in the rejection, the discriminator threshold is not raised, to not loose further signal efficiency.

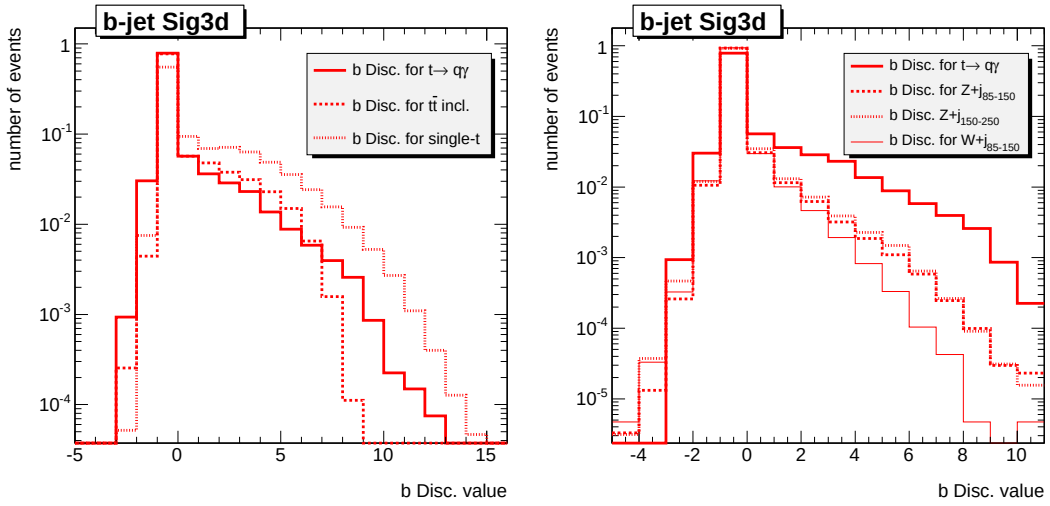

Figure 3.65: The distribution for the b-tagging discriminator variable, for signal and the relevant background (that will survive even in the last cut) and for signal and W/Z+jets, that have large cross section and are annoying till the last selections. No optimal choices are offered, thus the $Disc > 2.0$ is established.

When the request of the b-jet is added to the selection, efficiencies are lowered to the values in Tab. 3.28. In this way, di-boson processes are hardly suppressed and exit the game, while processes with one or more b-jet are reduced by a factor ~ 3 that, as estimated for the previous analysis, correspond to the 'raw' efficiency for a single tagging. Boson+jet processes have no bottom jets, enters the sample only thanks to the mistagging so they are reduced of large factors as well: on the other hand, the W+jets process can still play a role.

Selection of the final state light-jets

The light-jet selection, applied to the signal and to all backgrounds considered, produces the p_T distributions showed in Fig. 3.66 (left). No significant differences exist in the signal behaviour (upper plot) from the $t \to qZ$ analysis, demonstrating that a common cut selection optimization works fine. As discussed below, the weak purity visible here is due to the mis-interpretation of b-jets as light ones, and it is the price to be payed to not prefer the selection of a u quark from the FCNC decay rather than a c.

The stacked plot shows that di-boson sources are not visible on top of the huge W+jets production. The other contributions come from single-top and from top pair production, here inclusively considered. Similarly to the previous case, asking for the presence of at least one b-jet is instrumental to strongly mitigate the impact of boson plus jet processes. On the other hand, single and pair top production will be enhanced.

It is interesting to observe the distribution of the light-jet multiplicity, *i.e.* the fraction of events containing no jets or a specific number of jets. The comparison with signal of the major background is presented in Fig. 3.66 right), and a process with a very high rate as the

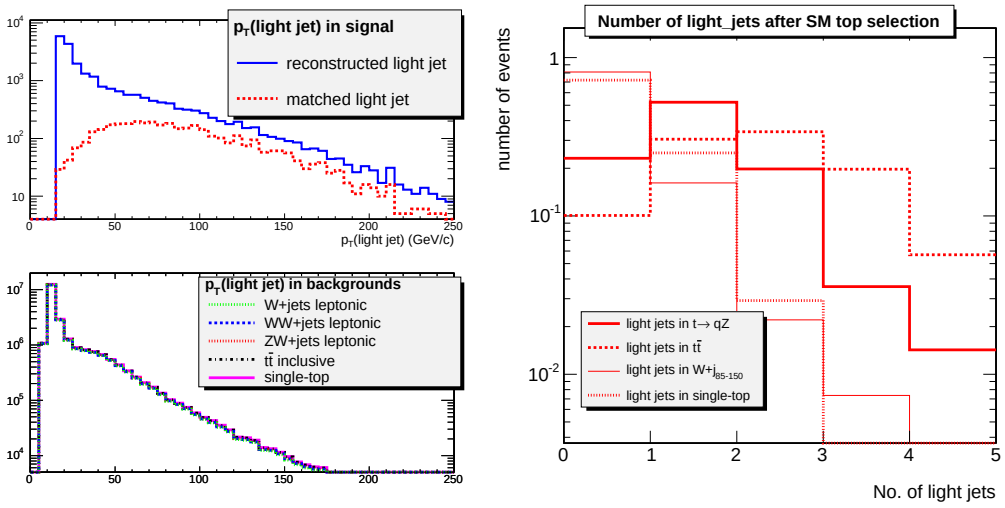

Figure 3.66: Left: The distribution of the transverse momentum of light-quark jet candidates in both the signal sample and the different background samples. In the lower plot, where background sources are stacked, vertical scale represents the number of events after $10\,\mathrm{fb}^{-1}$. As in the $t \to qZ$ case, p_T threshold is considered at $60\,\mathrm{GeV}/c$. Right: the multiplicity of the selected light-jets.

W+jets is also observed.

3.4.4 Constraints on the mass of the top with SM decay

As for the previous decay channel, the top quark following the Standard Model decay is identified by its transverse mass from chosen W boson and b-jet. A quite similar shape can be appreciated from Fig. 3.67(upper), where $M_T(bW)$ is again compared with the transverse mass of a jet matching the b quark and a lepton matching the lepton from the generated W. The good agreement with these matched variables enforces the statement that the correct object has been reconstructed. In each event, only one such object selected in this way is asked for.

For what concern the background, the situation is remarkably different from the $t \to qZ$. No powerful suppression requirements as the presence for a third lepton is possible to ask here, thus the contribution of background processes featuring a top with Wb final state is important. In addition, the large transverse mass window drains a part of the boson plus jet sample, that can pass the selection when the light-jet is mistagged as a b one. Due to its giant production cross section, the process appears as completely dominating the selection, but it will readily rejected when the additional (light) jet and the hard photon are asked for. Contributions from $Zb\bar{b}$, although not showed here, continue to be sizable.

An estimation of the significance s_x for this selection, plotted as a function of the mass upper and lower thresholds in Fig. 3.68(left), shows that even in this case statistical properties are improved if the tail in the distribution having $M_T(bW) > 220\,\mathrm{GeV}/c^2$ is cut away. Robustness around this cut is again satisfactory: it can be evaluated that, if the threshold

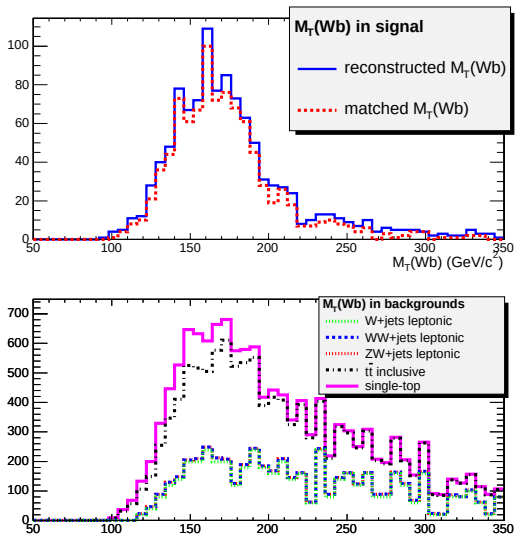

Figure 3.67: The invariant transverse mass distribution for the top with the standard decay, in the signal and the different background samples. The comparison with a generated b quark close to the reconstructed jet and a matched W demonstrates a nice agreement. The background samples are normalized assuming an integrated luminosity of L=10fb^{-1}.

Table 3.29: The efficiency for the (optimized) selection of one top following the Standard Model decay. All the background sources considered are included. Errors (all in percent) are statistical only.

Selection	Signal $t \to q\gamma$	$t\bar{t} \to l + X$	single-t	$Zb\bar{b}$	$W + j_{85-150}$
one $t \to Wb$	6.37 ± 0.20	1.89 ± 0.07	0.1070 ± 0.0010	0.509 ± 0.012	0.0354 ± 0.0030
Selection	$ZZ \to 4l$	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one $t \to Wb$	0.0134 ± 0.0020	0.055 ± 0.004	0.083 ± 0.005	0.0047 ± 0.0017	0.029 ± 0.03

value is shifted by about 10%, the decrease of s_x induces a relative variation in the signal efficiency around 2%.

When the transverse momentum for the reconstructed Standard Model top quark is calculated, distributions exhibit a quite different behaviour. Figure 3.68(right) compares the $p_T(Wb)$ distributions for signal, $t\bar{t}$ inclusive, single-top in t-channel and $Zb\bar{b}$ also. While processes where top quarks are pair-produced show a wide spectrum in the top boost (and $Zb\bar{b}$ closely emulates this feature), single top production is peaked around 40-50 GeV/c. This may ascribed to the fact that this top does not recoils with another top but against a much lighter object as a light quark: therefore, this latter takes a reduced transverse momentum, and the top results softer. The effect nicely agrees with what observed for the W from top. Since the different hardness of this object was already exploited there, no cuts on the $p_T(Wb)$ appear necessary.

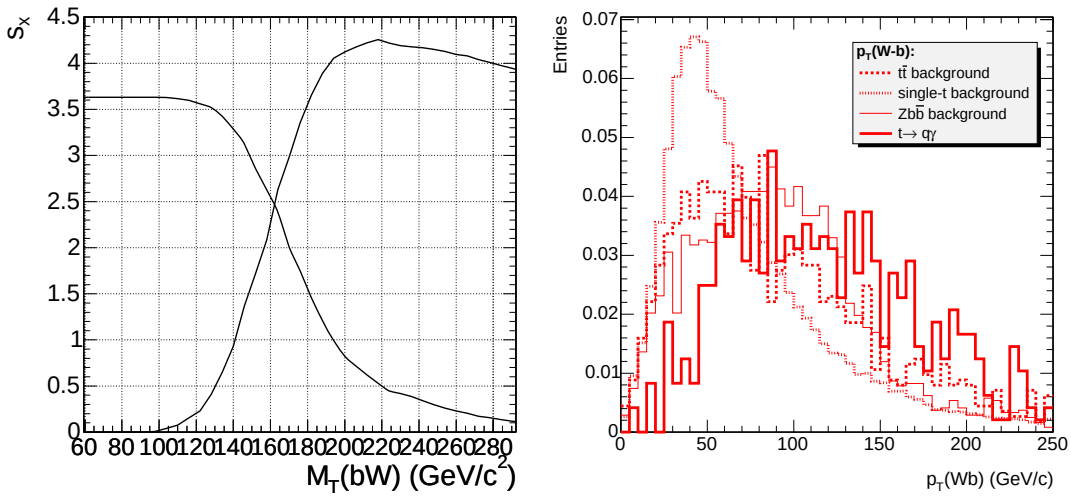

Figure 3.68: Left: The statistical significance of the selection in a $M_T(Wb)$ window, in function of the lower (left) and upper (right) mass cuts. A large cut as $M_T(Wb)$ <220 GeV/c^2 seems the best choice. Right: The transverse momentum of the reconstructed top with $t \to Wb$ decay, along with relevant processes entering the invariant mass selection. The shapes are normalized to the same number of events and preselection of the $p_T(W)$ is not applied here.

3.4.5 Constraints on the top with FCNC decay

As in the previous case, the identification of the top decaying through the Standard Model way is the most important flag for the anomalous decay reconstruction. The selected photon and the light-jet are constrained in an invariant mass bound. Again, the combination whose invariant mass lies closest to the nominal top mass is assumed to be the correct one.

A distribution as the one in Fig. 3.69(upper) is obtained. The overlapped histogram (dashed line) is the invariant mass done with a photon and a light-jet close to a generated γ and u/c quark, under the condition that the standard top has been reconstructed. Though statistic is limited, the agreement between 150 and 200 GeV/c^2 is good and room for the combinatoric is left only in the tails.

This reduced impact of combinatorics can be ascribed to the presence of few mass constraints, so avoiding to accumulate non-genuine events under the mass peak. The composed function $f_{B+S}(x)$ (Eqq. 3.6 and 3.7) has been adapted again in the $(x_L, x_H) = (100, 300)$ GeV/c^2 range, in order to maximize the statistic power of the calculation. The result of the operation is presented in Fig. 3.69(lower), where the signal has the error bars obtained from a quadratic sum of Poissonian uncertainties, and both the total and the background function are overlapped. The fitted value for the Gaussian and the background fraction are indicated in the box.

Once the fit has converged, the number of signal and combinatorial background is found again by an analytic integration of the fitted function. The fitted number of signal events turns out to be $N_S = 170 \pm 19$, that corresponds to an efficiency $\epsilon_S = (4.0 \pm 0.4)\%$. This is

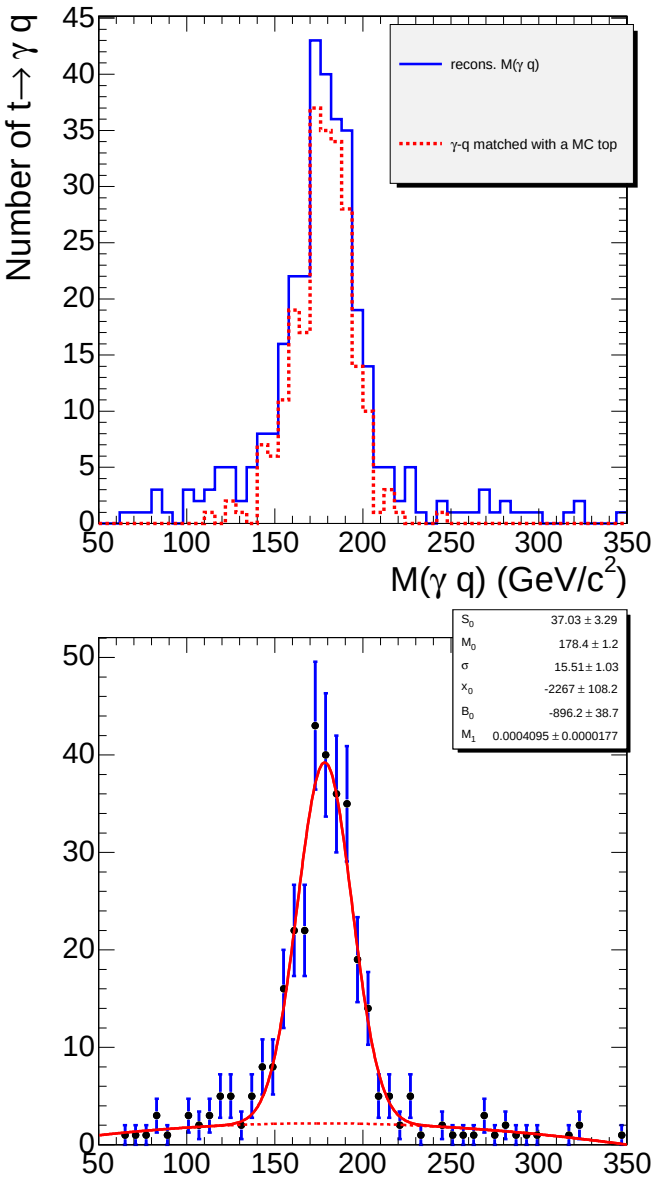

Figure 3.69: Upper: The reconstructed invariant mass of the photon and the light quark (solid), compared with the same object when matched to the generated particles (dashed). Each bin contains 6 GeV. Lower: The result of the fit on the top FCNC signal distribution with a Gaussian (for the 'genuine' signal) summed to a linear plus exponential (for the combinatorial background) function. Total function (solid) and background (dotted) are superimposed and resulting fitted parameters are in the table.

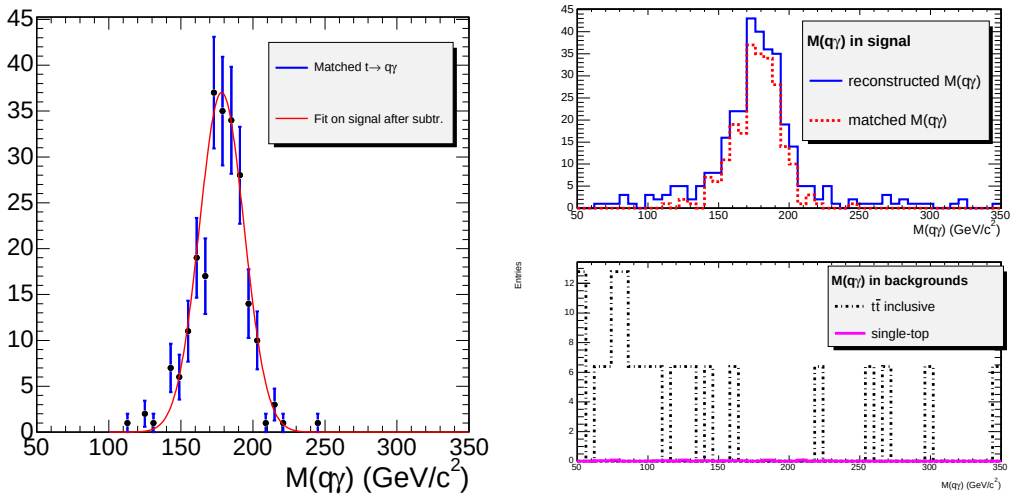

Figure 3.70: Left: The result of the Gaussian part of fit on the top FCNC signal distribution, compared with the histogram of the matched object. The agreement is good and demonstrates that the aim to extract the lineshape of the top quark from its decay on top of a rich background has been reached. Right: The final selection for a $t \to q\gamma$, that requests the hard photon to the sample, when applied to signal and all background. The single-top is a tiny contribution at the bottom of the $t\bar{t} \to l + X$ inclusive production. Vertical scale corresponds to the counting scaling to $10\,\mathrm{fb}^{-1}$.

assumed as the final estimate of the selection efficiency for the $t \to q\gamma$ signal.

When the fitted Gaussian function is compared with the 'matched' invariant mass distribution – that is performed using a light-jet closer ($\Delta R < 0.3$) to the generated one and the matching condition for the photon and the standard top decay is ensured – the result in Fig. 3.70(left) appears. Although here the efficiency is lower and error bars larger, the good quality of the final reconstruction is again demonstrated.

The comparison with the other background sources, displayed in Fig. 3.70(right), shows that the additional requests for the hard jet and the high energy photon success in suppress several background sources but, in addition to the inclusive $t\bar{t}$ production and the single-top, also a part of the Z+jets and $Z + b\bar{b}$ gives a contribution. The two most relevant processes are shown stacked and the single top appears as a small ripple at the bottom of the top pair production. As in the previous analysis, fractional numbers come from the rescaling of the rate to the $10\,\mathrm{fb}^{-1}$ integrated luminosity.

It is worth to note that the amount of surviving event is measured with the only aim to evaluate the efficiency of this last cut. In order to improve the statistical uncertainty on background efficiency, the cut on the photon transverse momentum has been relaxed down to the trigger threshold and efficiency recalculated. This leads to an enhancement of the surviving events number, so lowering the counting error, without significantly affecting the light or b-jets selection efficiencies. Such efficiencies are then factorized with that from the optimized $p_T(\gamma)$ cut.

The efficiencies for the final selection of the FCNC channel are contained in 3.30, for the

Table 3.30: The efficiency for the last selection in the analysis for $t\bar{t} \to (\gamma q)(W\bar{b})$ signal. The only surviving processes are the one producing a real top quark, namely the single and the associated top production. The subtraction of combinatorial background from the signal distribution leads to a $(4.0 \pm 0.4)\%$ signal efficiency.

Selection	Signal	$t\bar{t} \to l + X$	single-t	$Zb\bar{b}$	$W + j_{85-150}$
one $t \to Wb$ and one $t \to \gamma q$	5.52 ± 0.17	$(1.71 \pm 0.14) \cdot 10^{-3}$	$(8.7 \pm 4.5) \cdot 10^{-5}$	0	0
$100 < M(q\gamma) < 250$	5.15 ± 0.16	$(1.71 \pm 0.14) \cdot 10^{-3}$	$(6.7 \pm 3.4) \cdot 10^{-5}$	0	0

Selection	$ZZ \to 4l$	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
one $t \to Wb$ and one $t \to \gamma q$	0	$\sim 10^{-4}$	$(1.7 \pm 0.6) \cdot 10^{-3}$	0	0

signal and all the simulated background sources. The zero values here correspond to $< 1 \cdot 10^{-5}$ for the $W + j$ process, $< 2 \cdot 10^{-5}$ for the $ZZ \to 4l$ process, $< 3 \cdot 10^{-5}$ for the $WZ \to 3l$ and $< 3 \cdot 10^{-5}$ for the $WW \to 2l$.

The suppression factors for the $W + b\bar{b}$ production and single-top tW-channel have been found to be respectively larger than $2 \cdot 10^5$ and $5 \cdot 10^5$, therefore they will not play any role in the final selection.

3.4.6 Summary of the $t \to q\gamma$ analysis

Revealing the FCNC signal $t\bar{t} \to (\gamma q)(W\bar{b}) \to (\gamma q)(l\nu b)$ has presented some more challenges than the previous case, because additional background sources enter the game and less powerful selection cuts are at hand. Nevertheless, a signal selection efficiency of the order of percent seems feasible, with a surviving background event number of the order of unity at $10\,\mathrm{fb}^{-1}$.

The main selection requirements for the this channel are summarized below; each item corresponds to a row in Table. 3.31:

1. the "single photon" stream has turned out to produce the best background rejection, so the event is triggered by a hard photon;

2. offline, each event must contain either a isolated $\mu^{\pm}$ (with $p_T > 20\,\mathrm{GeV}/c$) or an $e^{\pm}$ (with $p_T > 30\,\mathrm{GeV}/c$), plus more than 25 GeV of missing transverse energy. The combination of the lepton candidate and the missing transverse energy must have a transverse invariant mass lower than 120 GeV and a transverse momentum greater than $65\,\mathrm{GeV}/c$ for $p_T(e\text{-}E_T^{miss})$ or $50\,\mathrm{GeV}/c$ for $p_T(\mu\text{-}E_T^{miss})$. Events with more than one good W candidate are rejected;

3. each event must contain exactly one jet (with $p_T > 50\,\mathrm{GeV}/c$) which is compatible with coming from a b quark and is selected with a discriminator value greater than 2;

4. when a transverse invariant mass is produced combining of the b-jet with the W candidate, the upper limit $M_T(bW)$ <220 GeV/c^2 is again the best;

5. the isolated photon selected from the trigger must satisfy the p_T >90 GeV/c off-line cut. One jet incompatible with coming from a b quark is further requested, that has to be harder than 60 GeV/c. The combination of these objects must have an invariant mass in the range 100< $M(q\gamma)$ <250 GeV/c^2 .

Each cut has been tuned with the aim to maximize the statistical significance of the result that, as it will be proved in the next chapter, is proportional to the $N_S/\sqrt{N_B}$ ratio.

Table 3.31: The efficiencies (in percent) for the most important selections cuts in the $t \to q\gamma$ analysis. Zero values mean that no event from the chosen Monte Carlo sample is expected to survive; statistical upper limits are detailed in the text.

Selection	Signal $t \to q\gamma$	$t\bar{t} \to l + X$	single-t	$Zb\bar{b}$	$W + j_{85-150}$
HLT Trigger	95.74 ± 1.2	26.0 ± 0.4	2.22 ± 0.04	80.1 ± 0.3	37.03 ± 0.14
one good $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	24.2 ± 0.5	8.59 ± 0.14	0.60 ± 0.06	8.64 ± 0.07	5.99 ± 0.04
one good $W \to l\nu$ and one b-jet	6.8 ± 0.3	3.17 ± 0.09	0.222 ± 0.002	1.22 ± 0.02	0.089 ± 0.005
one $t \to Wb$	6.37 ± 0.20	1.89 ± 0.07	0.1070 ± 0.0010	0.509 ± 0.012	0.0354 ± 0.0030
$100 < M(q\gamma) < 250$	5.15 ± 0.16	$(1.71 \pm 0.14) \cdot 10^{-3}$	$(6.7 \pm 3.4) \cdot 10^{-5}$	0	0

Selection	$ZZ \to 4l$	$Z + j_{85-150}$	$Z + j_{150-300}$	$WZ \to 3l$	$WW \to 2l$
HLT Trigger	70.9 ± 0.3	57.76 ± 0.18	64.0 ± 0.2	47.9 ± 0.3	67.3 ± 0.3
one good $W \to l\nu$ (with $M_T(l\text{-}E_T^{miss})$ cut)	2.88 ± 0.03	4.72 ± 0.03	4.77 ± 0.04	4.53 ± 0.05	7.87 ± 0.05
one good $W \to l\nu$ and one b-jet	0.024 ± 0.003	0.097 ± 0.005	0.217 ± 0.009	$(9.6 \pm 2.3) \cdot 10^{-3}$	0.183 ± 0.008
one $t \to Wb$	0.0134 ± 0.0020	0.055 ± 0.004	0.083 ± 0.005	0.0047 ± 0.0017	0.029 ± 0.03
$100 < M(q\gamma) < 250$	0	$\sim 10^{-4}$	$(1.7 \pm 0.6) \cdot 10^{-3}$	0	0

3.5 Other background processes in the selection windows

The last section of this chapter is intended to address – in a mostly qualitative way – the processes that in principle could have some effect on the signal extraction, and have not been included in the simulation. Such processes can be either the Standard Model of the new physics ones.

In fact, in the Standard Model there are many processes with a sizable cross section that can largely share some of the distinctive features of the signal signature. On the other hand, such a signature includes very strong cuts as well, and at the end of selection procedure only very few background sources are supposed to survive. Simulation of these processes has not be undertaken here, because the large cross section would have required long computing time and large disk space, or because the multi-jet topology has not allowed to achieve a good precision in the cross section estimate – at least not yet comparable with the simulated channels. The processes with the largest production rate, as QCD background, exclusive $t\bar{t}$ production channels or $Z + bj$ are briefly discussed in the following.

In addition, the possibility that in the selected phase space there is some window for new physics different from the searched FCNC, is worth to be explored. Among the channels for which CMS has been designed, *i.e.* Higgs production and SUSY, final states have been considered that could resemble the signal signature. The possibility to detect a different FCNC channel (as the three-body decays mentioned in Sec. 1.3) and different decay modes of the signal are included in this estimation.

In most of the cases, reasonable motivations will be issued to confirm that the impact of these processes is totally negligible. Therefore, not having included them in the analysis from the beginning makes no big hurt to the whole study.

3.5.1 Analysis of the multi-jet background sample

A huge QCD multi-jet production is known to be a key issue for the collisions at LHC, and most of the analyses have to face the challenges it offers. Possible impact of the jet production process have been suggested in Sec. 3.1.2, where the rates have been explicitly quoted as a function of $\hat{p}_T$ of the outgoing parton. From these values, it is evident that a detailed simulation of such background would need samples of events of more than 10^9 or 10^8 at the first integrated inverse femtobarns. Computation time and storage space put strong limits to this operation, thus different approaches are mandatory. Factorizing the signal selection in a set of partial independent cuts is a common approach, and it will be adopted here.

The factorization method

The purpose of this qualitative study is to devise a strategy to extract a rejection factor for the QCD background processes and verify whether it is possible for some events to pass all

the selection requirements. As explained earlier, chances to emulate the signal may exist only for the $t \to q\gamma$ selection, so it will be the only addressed here.

To accomplish the task, the full procedure is separated into a set of uncorrelated selection requirements and the absolute efficiency for each of them is computed. The rejection factor for the global selection will then be the product of the efficiencies for each independent requirement. The selection requirements for which efficiencies have been separately evaluated are given below, along with a conventional denomination:

1. each event must contain two jets with $p_T > 50$ GeV/c, without the application of b-tagging techniques and a missing transverse energy which satisfies $E_T^{miss} > 35\,\mathrm{GeV}$ (*2 hard jets*);

2. each event must contain an isolated photon which satisfies $p_T > 90\,\mathrm{GeV}/c$. In fact, this requirement could be anti-correlated with the item 1 above, because the jet could be identified as the photon and vice versa. The assumption that they are actually uncorrelated must therefore be considered to be strongly conservative (*hard* γ);

3. each event must contain a good lepton candidate. The real lepton from the W decay will be naturally correlated with the missing energy from the neutrino, but it has been observed that the majority of the reconstructed leptons in the multi-jet background samples are actually mis-identified hadrons. It is therefore a reasonable assumption that this selection requirement is uncorrelated with the item 1 above (*hard lepton*);

The b-tagging selection is not applied to the surviving sample, in order to retain events for further study. Under the reasonable assumption that the vast majority of jets in this background sample will not have been produced by b-quarks, the events which survive the b-tagging procedure should be predominantly mis-tagged light-quark jets. In this case, the mis-tagging efficiency given in Sec. 3.2.3 can be used as the uncorrelated efficiency for this selection requirement (labeled as *b-tag*) in Tab. 3.32. The dominant contribution in the QCD multi-jet background sample comes from the light quark and gluon-initiated jets ($\sim 70\%$) that can in turn split into light flavours. As Sec. 3.2.3 pointed out, this efficiency can conservatively considered to be around 1%.

The efficiencies for these uncorrelated selection requirements are given in Tab. 3.32, along with the overall combined efficiency and the number of expected events.

Conclusion on QCD impact

The overall effect of these selection requirements is to reduce the very large multi-jet QCD background contribution to approximately 24 events, assuming an integrated luminosity of $10\,\mathrm{fb}^{-1}$. The surviving events occur mainly in the range ($100 < \hat{p}_T < 500$) GeV/c. It is possible to argue that these events will fail the mass constraints placed on the W and

Table 3.32: The absolute efficiencies for the four independent selection requirements described in the text. The values are quoted both for the multi-jet QCD background sample and the SM $t\bar{t}$ sample. The last two columns contain the combined rejection factor and the number of expected events from a dataset with an integrated luminosity of $10\,\mathrm{fb}^{-1}$.

Cut	*2 hard jets*	*hard* γ	*hard lepton*	*b-tag*	*Total Eff.*	*Expected evs.*
QCD $50< \hat{p}_T <80\,\mathrm{GeV}/c$	0.1%	0.04%	0.03%	1%	$1.2 \cdot 10^{-12}$	< 0.1
QCD $80< \hat{p}_T <120\,\mathrm{GeV}/c$	0.8%	0.2%	0.1%	1%	$1.6 \cdot 10^{-10}$	4.9
QCD $120< \hat{p}_T <170\,\mathrm{GeV}/c$	2.3%	0.3%	0.1%	1%	$6.9 \cdot 10^{-10}$	3.4
QCD $170< \hat{p}_T <230\,\mathrm{GeV}/c$	5.7%	0.6%	0.2%	1%	$6.8 \cdot 10^{-9}$	6.9
QCD $230< \hat{p}_T <300\,\mathrm{GeV}/c$	9.3%	0.6%	0.2%	1%	$1.1 \cdot 10^{-8}$	2.7
QCD $300< \hat{p}_T <380\,\mathrm{GeV}/c$	13.3%	0.7%	0.6%	1%	$5.6 \cdot 10^{-8}$	4.3
QCD $380< \hat{p}_T <470\,\mathrm{GeV}/c$	15.2%	0.6%	0.6%	1%	$5.5 \cdot 10^{-8}$	1.0
QCD $470< \hat{p}_T <600\,\mathrm{GeV}/c$	16.2%	0.5%	0.5%	1%	$4.0 \cdot 10^{-8}$	0.3
QCD $600< \hat{p}_T <800\,\mathrm{GeV}/c$	17.4%	0.5%	0.7%	1%	$6.1 \cdot 10^{-8}$	0.2
QCD $800< \hat{p}_T <1000\,\mathrm{GeV}/c$	17.3%	0.4%	1.3%	1%	$9.0 \cdot 10^{-7}$	0.3
$t\bar{t}$ *inclusive*	23%	0.8%	30%	37%	$2.0 \cdot 10^{-4}$	1702

top candidates and therefore the contribution from this source of background may well be significantly smaller.

The additional reduction factor from applying mass constraints can be estimated using the SM $t\bar{t}$ sample, as this has higher remaining statistics after the application of the uncorrelated selection requirements than the multi-jet sample. As there will clearly be a correlation between the missing energy requirement and the lepton selection in this case, only a lower limit on the final efficiency can be estimated. A comparison of the results for the $t\bar{t}$ background sample given in Tab. 3.32 and the number of $t\bar{t}$ events that remain in the full analysis implies that the final mass constraints have a rejection power better than 1.9%. Assuming that this is also true for the multi-jet QCD background sample (a rather pessimistic hypothesis), then this suggests that only $\sim 25 \times 1.9\% \simeq 0.5$ events will remain after the application of all cuts. It is therefore reasonable to neglect the multi-jet QCD sample as a source of background for this analysis.

3.5.2 The impact of additional Standard Model processes

The effect of background processes is further scrutinized here, by mentioning few other final states that were not possible to generate by PYTHIA, or that have been excluded from the beginning without motivations.

Multi-jet $t\bar{t}$ production and other decay channels

Calculation of $t\bar{t}$ production with the matrix element method succeeded in calculating processes that are beyond the tree-level production, as the $t\bar{t} + nj$ and $t\bar{t} + n\gamma$. The $t\bar{t} + nj$ diagrams are typically calculated with the ALPGEN program and have a cross section of some tens of pb (for not-b-jets, less or equal than 4). If some of these events meet all the signal selection requirements, they can increase the number of surviving $t\bar{t}$ and appear as an additional contribution under the $M(qZ/\gamma)$ distribution. The next chapter will point out that there is not need to account separately for these effects, as they will be included automatically in the estimation of background from data. Proper control regions will be defined that are sensitive to these additional objects, so the impact of the higher order effects will be directly measured from them, and these processes will anymore be a issue here.

Similar arguments hold for the $t\bar{t}+\gamma$ production. These processes will be treated in some more details when the effect of initial and final state radiation will be discussed.

The occurrence of $t\bar{t}$ production with any production of b-jets could be issued only in the case of a double $t \to d/sW$ decay channel. Since the branching ratio of a single light decay is of the order of 10^{-3}, the chance is eagerly remote and this channel, though possible, is never taken into account.

It is important to observe that in the $t \to qZ$ analysis, a very specific decay channel has been considered, *i.e* the $t\bar{t} \to (bl\nu)(\bar{b}l\nu)$ (with $l = e, \mu, \tau$), simply discarding all the others.

The analysis starts by choosing a di-lepton trigger, so there is no danger to include the purely hadronic top pair decay in the selection. An issue may rise only from the semileptonic decay, that could have some chance to produce a signature similar to the signal. The effect occurs if a further lepton from semileptonic b quark decay is mis-identified as a lepton from W/Z. By comparing the efficiencies to detect three and two leptons in the $t\bar{t} \to qZ$ sample, one could argue that this can happens about $(0.90 \pm 0.08)\%$ of times. Even if enhanced by the $BR(W \to q\bar{q}')/BR(W \to l\nu) \simeq 2$, the rejection factor is undoubtely high, thus the exclusion of this decaying mode is no limiting.

The $Z + bj$ production

Recent next-to-leading-order calculations [134] have shown that the production of a Z boson with two jets, only one of which contain a b quark, is larger than that with two b quark jets at both Tevatron and LHC. The total cross section for $bq \to Zbg$ at NLO amounts to 352 pb and the inclusive cross section for the $Z + bj + X$ (with $p_T(j) > 15\,\mathrm{GeV}/c$ and $|\eta(j)| < 2.5$) is $510^{+84+44+25}_{-58-35-25}$ pb, with uncertainties from the variation of the renormalization scale, the factorization scale and the parton distribution functions, respectively. The kinematic range of the two jets may well accommodate within the signal selection cuts, and the final state could be faked if a hard lepton is picked from a jet. With a conservative assumption, the probability for such a faking is taken again as the largest efficiency for the two and three leptons detection in Z+jet. From the Tab. 3.20, it turns out to be $(18 \pm 3)\%$. A further reduction factor arises from the transverse momentum cut in both the light and the b-tagged jets, that makes only the Z with $p_T(Z) < 50 \div 60\,\mathrm{GeV}/c$ to be relevant in the final selection. From Ref. [134] one could argue that only a portion around $\sim 15\%$ of the total cross section is taken: as a consequence, only a $Z + j + b\text{-}jet + X$ production cross section about 14 pb has to be taken, that is over a factor 50 smaller than the $Z + b\bar{b}$ production. In all, such Standard Model contribution can be fairly neglected.

3.5.3 The impact of other new physics processes

In addition to the well-known Standard Model processes, there could be several processes that can overlap to the searched signal, that are still to be observed but should be at LHC. Here the sensitivity is explored firstly for decay modes in the signal, different from the addressed ones, along with other FCNC possibilities. A brief discussion follows, aiming to infer if some SUSY or Higgs processes that LHC might produce have some impact on the signal extraction.

Other decay channels in the signals

The trigger menu proposed for the two channels (di-lepton for the $t \to Zq$ and single-photon for the $t \to \gamma q$) allows to efficiently reject the hadronic Z and W channels decay. The effect has been confirmed by the generation of a $\sim$ 1000 evs. subsample for the $t \to Z(\to q\bar{q})q$

signal, where a trigger efficiency of 0.4% has found (with the two leptons pairs outside the Z peak). Two further classes of signal events could be an issue:

- the hadronic decay of the W boson from the $t \to bW$. In the $t \to Zq$ analysis, the detection of the third lepton has found to occur in about 3% of the $W \to q\bar{q}'$ decay and can be reduced by a factor 2 by tightening the electron isolation cut from 0.07 to 0.04 (that has found to induce not a big hurt on the electron efficiency). The effect has not be tested on the $t \to \gamma q$ but a smaller magnitude can be argued, as the p_T cut on the lepton from W is harder;

- the tau decay of the W boson from the $t \to bW$. Instead to check this contamination by re-generating a sample of $t \to bW \to b(\tau\nu)$ in the signal, the fraction of events containing at least a lepton matched with a $e^{\pm}/\mu^{\pm}$ coming from $\tau^{\pm}$, in the $t\bar{t}$ sample after the leptonic cuts, is estimated. In the analysis for the $t \to Zq$ search a small 0.1% of residual is found after the Z request, while the $t \to \gamma q$ study can be sensitive to about 3% of the W decays with tau. On the other hand, such contamination is present both in the signal and in the $t\bar{t}$ background, that has demonstrated to be the larger one. For a 3% increase in N_S and N_B, the significance (that will be defined in the next chapter), would be affected by only a 0.8% shift. A corresponding variation in the FCNC branching ratio estimate will be induced, that has not any relevance.

As a consequence, the system of selections devised below demonstrates that the analysis is sensitive only to the generated FCNC signal and other decay modes have no sizable impact.

A different approach has to be taken for the three-body $t \to cV_iV_j$ decays, that could be possible in different models with specific EWSB parameters. It is important to observe that the same selection for the FCNC two-body decay can be adapted (with few minor modifications) to detect signals as $t \to cZZ$, $t \to cWZ$ or even more exotic like $t \to cW/Z\gamma$. In Sec. 1.3 it was mentioned how models as 2HDM (Type-III) predict branching ratios as large as $10^{-3} \div 10^{-2}$ for optimal choices of $\tan\beta$. Though the analysis for these double-boson signals is not performed here, it is easy to guess that the sensitivity can be pushed even further than the $t \to Vq$ channel, since the multi-leptonic signature will have a very low background. On the other hand, overlaps between the two signals can be minimized by identifying a narrow peak in the reconstructed top mass, with procedures similar to those designed above. In all, the two kinds of channels can be discriminating just by counting the number of W or Z boson in the final state.

Higgs associated production

The Higgs associated production with a $t\bar{t}$ or $b\bar{b}$ production has been proposed in a wide range of analysis, especially in order to detect the dominant decay channel $H \to b\bar{b}$, otherwise impossible to observe, as a result of the combination of an overwhelming QCD cross section for $b\bar{b}$ production and the inability to reconstruct the Higgs mass very precisely.

The $t\bar{t}H$ production could share several features with the signal signature, depending from the final state in which Higgs boson decays. The total production cross section (NLO) ranges from 0.747 pb to 0.525 pb for Higgs boson mass from 115 GeV/c^2 to 130 GeV/c^2 [135].

The $t\bar{t}H \rightarrow (Wb)(W\bar{b})(H \rightarrow b\bar{b})$ is one of the most promising channel for the Higgs observation via the associated top production and it may have a cross section up to about 0.5 pb [136]. With respect to the signal, not only it would be scaled three-times by the b-tagging efficiency, but the lepton mistagging should be very large to fake the signal. As a consequence, this decay has no chance to play any role in both the addressed signals.

In principle, sizable contributions could be delivered only by $t\bar{t}H \rightarrow (Wb)(W\bar{b})(H \rightarrow WW/ZZ)$ or even $t\bar{t}H \rightarrow (Wb)(W\bar{b})(H \rightarrow \tau\bar{\tau})$. If one would ascribe the leptonic contribution in the signature to the W/Z decays from the Higgs, it turns out that cross sections as small as $0.1 \div 0.2$ pb would be produced. Larger values may rise if one assumes the $t\bar{t}$ to affect the analysis in the way described in this chapter, and the hadronic decays of W/Z from Higgs as a source of additional light jets. Even inserting the largest branching ratios for $H \rightarrow WW$ and $H \rightarrow ZZ$ and considering that the probability to mis-identify a light jets from Higgs as a one from top FCNC decay is 100%, a contribution about 1 pb (with acceptance cuts) is expected. Finally, considerations made above for the top decay channels with a tau hold even for the $H \rightarrow \tau\bar{\tau}$.

Therefore, the whole signal selection procedure is not expected to be significantly sensitive to any Higgs decay.

Chapter 4

Sensitivity to FCNC in the $t \to qZ/\gamma$ channels

The aim of this chapter is to transform the results of the selections devised in the previous study into a realistic estimate for the CMS sensitivity to top FCNC. In the first section, a detailed study of systematic uncertainties on the analysis is outlined. The effect induced from both the limited knowledge of detector performance and the physics working under the simulated processes, are addressed. In many cases, the parameters that have been suggested to the whole CMS collaboration have been adopted, with different choices fully motivated. Equally, strategies for theoretical biases emulation are those exploited by the collaboration in the preparation of the Physics TDR, hence the results can be compared on a common ground.

Among the possible solutions to reduce the impact of the systematic effects, one of the most powerful is the estimation of background 'from data', *i.e.* from a well-defined region that will be accessible in the future data samples. A possible approach to accomplish this task is outlined in the subsequent section: therefore, the propagation of these uncertainties to the background is detailed there. At the end of the two sections, an estimate of the expected background and its cumulative uncertainty is available for both the decay analyses: when combined with the determined efficiency for the FCNC signal, the issue of statistical significance of the results comes up. After a general review of the concepts related to the significance definition, a specific significance region with the best statistical properties is identified. Finally, the evaluation of the significances for a discovery or confidence level for an upper limit is performed and dependencies from background level and systematic intensity will be underlined. An extrapolation is then performed to the high luminosity phase of LHC with a larger amount of pile-up, where the challenges from the different scenarios and the benefits from the larger statistic will be discussed.

4.1 The impact of systematic uncertainties

The present analysis, although in the low luminosity phase it could be considered still statistic limited, is supposed to enhance in precision along with accumulation of larger and larger data samples. Once the commissioning steps will be fully accomplished and a stable LHC run will be assessed, the statistical error should be limited only by time. Therefore, a control of systematic uncertainties is required and their impact on the analysis results have to be precisely determined.

The aim of this section is to address the most important sources of systematic effects, then emulate them in the generation step, or in the reconstruction phase, or either by changing parameters in the analysis in order to reproduce real data to the maximum possible extent. Since the background will be counted in specific 'control regions' – defined by a set of cuts as close as possible to the signal region – phenomena affecting the absolute background rate have no relevance here. Therefore, the impact of systematic effects is measured only for the variables that are chosen to mark the difference between the control region and the signal region, for the different situations of the two analyses. They are:

1. the lepton selection efficiency since, in the $t\bar{t} \to (Zq)(W\bar{b})$ analysis, the ask for a third lepton is a powerful tool to transform a region rich in background to a region rich in signal;

2. the missing energy resolution. As it will be demonstrated, a discrimination based on the E_T^{miss} distribution can disentangle very well the $t\bar{t}$ from the $Zb\bar{b}$ component in the $t\bar{t} \to (Zq)(W\bar{b})$ study;

3. the efficiency for b-jets and the multiplicities distribution, namely the fraction of single-tagged to multi-tagged events. The number of b-jets, as indicated in the previous chapter, is useful in both the analyses to better identify the different sources of background;

4. the efficiency for light-jets and their multiplicities distribution. The requirement for the hard jet is one of the last imposed to the sample, thus it is instrumental to define the signal region. Moreover, the number of light-jets will turn useful to disentangle the different backgrounds in the $t\bar{t} \to (\gamma q)(W\bar{b})$ analysis, that drains a richer set of contaminating processes in all but the last selection.

Since many of the parameters adopted to simulate the effects have a degree of approximation – if not just simply "reasonable"– some level of arbitrariness is intrinsic to the error estimates. On the other hand, exceedingly 'conservative' estimates have been avoided as possible and room enough is given to let the real data tell the final world.

4.1.1 The effect of instrumental systematic bias

Calibration of various sub-detectors and monitoring of their performance are critical issues for every kind of physical measurement. In the following, the instrumental effects mostly affecting the relevant efficiencies are addressed. Guidelines are basically dictated from studies detailed in Ref. [95].

Whenever possible, uncertainties are inferred even for the highest luminosity scenarios, in order to reach a good predictivity for the sensitivity in that region. As only real data will actually indicate how extrapolations can be done, these assumptions will be partially speculative.

All the fluctuations induced by systematic effects are estimated on samples "enriched" in background that, as detailed in the previous chapter, has been accomplished by broadening the Z mass window in the $t \to qZ$ analysis and softening the photon p_T lower cut, for the $t \to q\gamma$ channel. This trick is particularly important here, because the systematic effects, typically amounting to some percent, would not be visible on top of a larger statistic error. The impact of detector systematic uncertainties is detailed in tables from Tab. 4.1 to Tab. 4.4, for each sources for the most relevant background processes.

Luminosity uncertainty

As declared in Ref. [102], the design goal for the precision of the luminosity measurement at CMS is 5%, which is supposed to be achieved after 1 fb^{-1} of data has been collected. For integrated luminosities of 30 fb^{-1} or more, it is assumed that further improvement on the uncertainty can be achieved (*e.g.* via W,Z based luminosity measurements) and a 3% uncertainty is considered.

The uncertainties originating from the luminosity will only result in a fluctuation on the overall number of events, and no modifications of the efficiencies are to be expected. The absolute background rate will be estimated by counting in the control region properly optimized in the next section, so it is immaterial against any systematic effect. Therefore that procedure allows to avoid introducing the 3% error, that is one of its major benefit. Luminosity uncertainties are not an issue and they will not considered anymore.

Tracker and Muon chambers uncertainties

The main sources of systematic errors on the leptons identification and measure come from misalignments in the tracker and in the muon chamber, along with a small contribution from magnetic field uncertainties.

The degree of alignment of the two sub-detectors has been measured firstly by mechanical surveys during the sub-assembly phase, and has improved day-by-day by combined runs with cosmic rays. Tracks crossing the overlaps between modules in the tracker, for instance, are exploited to determine both the level of alignment and magnetic field. With the first p-p

collisions, clean physics samples (as $Z \to \mu\mu$) will be used to calibrate and align the detector. In one month of data taking, CMS will be able to collect $\mathcal{O}(10^6)$ $Z \to \mu\mu$ events so, thanks to this large statistic, this *in situ* calibration will deliver the ultimate precision to the alignment – well under $1\,\mu$m.

The tracker momentum scale and the muon momentum scale are considered as a fluctuation β_{TMS} and β_{MMS} on the transverse momenta, measured respectively by the tracker and the muon chambers. The effect on the analysis is reproduced by shifting the momentum scale alternatively in the negative and positive direction, separately for the two detectors. The fluctuations are assumed to include [137]:

- the effect of misalignments, in what CMS has called the "Long Term Scenario", that is appropriate after the first few fb^{-1}. For the low luminosity phase, it is supposed to be $0.0004/p_T[\,\text{GeV}/c\,]$ for both muons in the chamber and tracks in the tracker – that for the electrons is a quite conservative estimate – while for the high luminosity phase, where an high precision alignment is supposed to be reached, it is not introduced;
- the uncertainty in the magnetic field in the central region of CMS, affecting both the tracker and the muon spectrometer. It is assumed to contribute with $0.0003/p_T[\,\text{GeV}/c\,]$. Here it is supposed to remain constant during all the luminosity phases.

In total, it corresponds to an overall $0.0005/p_T[\,\text{GeV}/c\,]$ in the first phase and $0.0003/p_T[\,\text{GeV}/c\,]$ subsequently.

In all the backgrounds considered, the effect on the lepton detection does not exceed the 1%. Since leptons are involved in the procedure for the cleaning of all jets, a small impact is assessed also in the b and light-jet efficiency. A complete symmetry is found between the two shift directions.

In addition, a relative uncertainty of 1% is added as a fixed contribution to both muon detection efficiency and track detection efficiency (that propagates on electrons and jets efficiency). It is intended to reproduce the uncertainty on fraction of events in which the detection of a charged particle track or a muon has failed.

Electromagnetic energy scale

The use of single tracks (during the cosmic data or the first run phase) does not fully determine the energy scale of the ECAL, since deformed topologies satisfying the constraints are always possible. Therefore, calibration of energy scale is also obtained from a resonance with well-know mass decaying to a lepton, such as $Z \to ee$ and J/Ψ, depending from the specific energy domain.

The CMS goal is to achieve a ECAL global intercalibration better than 0.5%. High p_T isolated electrons are produced at a sufficient rate from $W \to e\nu$ and $Z \to ee$ at LHC (around 10 Hz and 1 Hz at low luminosity) and can be used to obtain intercalibration coefficients and

absolute electromagnetic energy scale. Neverthless, calibration with electrons is complicated by bremmstrahlung effects. A possible strategy is to apply loose cuts on bremmstrahlung to intercalibrate ECAL region with an uniform quantity of material in front, while selecting electrons with no or little bremmstrahlung to intercalibrate different calorimeter regions. Few hours data taking during low luminosity phase correspond to 11 million jet trigger events, thus they are sufficient to perform the intercalibration to a precision between 2% and 3% (depending on η).

No unique number is defined for the ECAL energy scale uncertainty, but they are dependent on the algorithms used to define electrons and photons. For the 'golden' electron presented in Sec. 2.3.1 (that basically correspond to those selected in the analysis) a 0.05% uncertainty is considered. It has reproduced by shifting the energy of the electrons and photons by a term $\beta_{EES} = 0.05\%$, that is conservatively considered constant for all the luminosity phases.

An impact from this effect is found only in the lepton and photon efficiency, and is lower than 0.5% in all the samples. This maximum value is thus assumed all along the different luminosity phases.

Jet energy scale

Since maximizing the mass resolution on the final state is not an aim of this analysis, the reconstruction of $t \to bW$ and $t \to Z/\gamma q$ final state does not need a very precise knowledge of the jet energy. On another hand, some delicate optimizations in the analysis rely on thresholds of mass distribution or transverse momentum: as a consequence, a systematic distortion in the energy measure is expected to have some impact on different efficiencies. Hereafter, 'energy scale' refers only to the transverse energy E_T of light quark jets and b-jets, as reconstructed from the CMS software and calibrated with the simulated event samples. Since ECAL can be calibrated from a variety of techniques up to a $0.4 \div 2.0\%$ precision, the bulk of the energy uncertainty comes from HCAL.

The jet energy scale $\alpha(E_T)$ is the factor that calibrates the energy collected from calorimeter towers, E_T^{RAW}, to the true energy of a jet, $E_T(j)$, as $E_T(j) = \alpha(E_T)E_T^{RAW}$. It depends on a variety of effects coming from detector (as non-linearities in the calorimeter response, cracks, leakage, electronic noise etc.) and physics (energy swept outside the jet cone due to the magnetic field or to gluon radiation at large angles, invisible energy, underlying event).

The determination of absolute energy scale for leptons and jets is one of the greatest challenges during the initial phase of the experiment. It requires the knowledge of a large numbers of detector parameters: the tracker alignment, the magnetic field map in the tracker volume, the tracker material distributions, the calorimeter calibration and the muon energy loss in the calorimeters.

At the LHC start-up the accuracy of the jet energy scale relies on single-particle test beam calibration and comparisons of simulated event samples with the detector response, and an

overall uncertainty of 15% in $\alpha(E_T)$ is expected.

In the first $1 \div 10$ fb^{-1} of data, the best estimates of the absolute jet energy scale are obtained with two methods:

1. hadronic W boson decays in $t\bar{t}$ production events [138]. They produce jets with a mean p_T around 50 GeV/c and uncertainty about 3% is originated mainly from pile-up. It is almost flat at high energy and increases when lowering the p_T threshold;

2. the gamma+jet calibration [114], that can set the energy scale calibration down to ~ 20 GeV/c.

Below 20 GeV, only the single-particle calibration methods apply and these will have an accuracy of 10%.

Therefore, the fluctuations of the energy scale are parametrized as $\alpha(E_T) = k(1 + \beta_{JES}(E_T))$, where $\beta_{JES}(E_T)$ takes account of the systematic fluctuations. Adopting the simple model proposed in Ref. [102], it consists in a linear falling between 5% (for $p_T < 20$ GeV/c) and 3% (for $p_T > 50$ GeV/c), *i.e.*:

$$\beta_{JES}(E_T) = \begin{cases} 5\% & \text{if } p_T < 20\,\text{GeV}/c \\ 5\% - 3\%(E_T[\text{GeV}] - 20)/30 & \text{if } 20 < p_T < 50\,\text{GeV}/c \\ 3\% & \text{if } p_T > 50\,\text{GeV}/c\,. \end{cases}$$

The effect of this uncertainty has been assessed by inserting this scale variation and measuring the corresponding changes in the event yield. The impact on the efficiency for b-jet is not an issue here, because the effect of asking for a b-tagged object will be directly included in the background observation from data. In fact, the definition of the control region emulates as close as possible the b-jet requirements for the signal region, in this way resulting insensitive to systematic effects related to these jets. Optimization cuts in the $M_T(Wb)$ presented in the previous chapter, that instead heavily rely on simulated event samples, are quite robust against energy fluctuations and even after a 10% variation the decrease in significance would be irrelevant.

The major effects for $\alpha(E_T)$ have to be expected when the light-jet is searched, since it marks the difference between the control and signal region, and the impact of inserting this object on the selection has to be measured from simulated event samples.

Results on the four samples with the major backgrounds show that this is one of the most important effects. The number of detected jets can vary in average about 15% in both the light and the b-jet selection. Part of this inefficiency propagates on the missing energy value (as detailed below) and results in a small broadening of the missing energy resolution. Modifications in the energy scale also induce a different occurrence of events with a single tagged jet, both light and b-tagged.

In order to emulate a better knowledge of the hadron energy scale, the study of this systematic effect is repeated with an overall reduction of the efficiency, namely a 3% when

$p_T(j) < 20\,\mathrm{GeV}/c$, 1.5% when $p_T(j) > 50\,\mathrm{GeV}/c$ and a corresponding scaling inbetween. With this choice, all effects seem to scale consistenly, indicating that a good control of the jet energy scale is instrumental in reducing the final systematic error.

Missing energy uncertainty

The definition of missing transverse energy adopted in the analysis (see Eq.2.3) shows two different contributions, for which different systematic bias are expected. The low p_T and unclustered energy component will be know up to a 10% precision following the first $1 \div 10$ fb^{-1} of data. On the other hand, high p_T jets have their own scale, correlated to the jet energy scale uncertainty of the high p_T jets.

Although the recommended treatment would be to apply separate uncertainties for the low and high energy components in the E_T^{miss}, the present analysis forces the jet to have a quite high p_T. Moreover, in the next section it will be shown that only the missing energy around the 150 GeV region is of relevance: therefore, E_T^{miss} can be assumed to closely follow the jet energy scale uncertainty.

As has been discussed in Sec. 3.2.2, the "muon+jet" method has been adopted for its best matching with the neutrino, so it inherits the energy uncertainties of these objects. In the present analysis, missing energy is calculated in this way only after that muon and jets have been rescaled for their uncertainties: consequently, fluctuations in E_T^{miss} have not to be simulated, because they are already included in its definition.

Uncertainties in b-tagging

The uncertainty in b-tagging refers to the experimental uncertainties on the number of b-tagged objects for a given sample of b-jets. Strategies to measure the tagging efficiency and its precision have been defined, for instance using a sample with a high purity in b quark content as the $t\bar{t}$ [139].

The result of these studies is that, in a b-tag working point efficiency of 50% and for a $10\,\mathrm{fb}^{-1}$ of integrated luminosity, the relative uncertainty on the b-efficiency measurement is expected to be about 7% in the barrel up to $40\,\mathrm{GeV}/c$, then smoothly decrease down to 4% up to about $100\,\mathrm{GeV}/c$ and remain costant for higher p_T. About a 1% greater uncertainty is expected in the endcap. Moreover, when more luminosity will be integrated an improves on b-tagging precision will naturally come, as efficiency should be directly measured from the abundant $Zb\bar{b}$ and $t\bar{t}$ events.

An uncertainty β_{BTAG} on b-tagging acts systematically in two senses, by exceeding in the number of tagged jets if β_{BTAG} is positive, or missing some of them if negative. The effect is here reproduced with a simple toy model, where a random number is sorted in a uniform distribution ranging from 0 to 1. If the number falls below $|\beta_{BTAG}|$, a candidate b-jet in the event changes its condition, being promoted to a tagged jet if $\beta_{BTAG} > 0$, being rejected to

the sea of un-tagged jets if $\beta_{BTAG} < 0$.

Since the jets with $p_T < 40\,\mathrm{GeV}/c$ are ruled out from the selection and the analysis is oriented to luminosity larger or equal than $10\,\mathrm{fb}^{-1}$, this uncertainty is simulated for $\beta_{BTAG} = \pm 5\%$ and $\beta_{BTAG} = \pm 2\%$, with the latter value proposed for luminosities greater than 10-30 fb^{-1}.

When such algorithms are applied to the analysis, important effects emerge, confirming the sensitivity of the whole selection procedure to details of the b-tagging algorithm. Firstly, impact depends from the direction in which tagging inefficiencies are pushed. When the number of b-jets is increased ($\beta_{BTAG} > 0$), the condition of a single-tagging is satisfied anyway and, even if the tagged jets are subtracted from the total amount, a light-jets somewhere is not difficult to find. When $\beta_{BTAG} < 0$, the number of tagged jets is lowered and the condition to find at least one b-jet with a high significance may be not satisfied. This situation would reject the whole event and gives no access to the light-jet selection. Such effect is confirmed in both the $t\bar{t}$ samples, in the $Zb\bar{b}$ (even if less clearly) and is enhanced in the single-top, where only a single b-jet is in principle available. A corresponding reduction is registered by introducing the 2% inefficiency, but the effect is still quite large. The fraction of b-single-tagged jet varies instead symmetrically in the two directions, and reads a value close to the β_{BTAG}. It indicates that the variations of b-jet rate occur mostly in the single-tag bin – that is the highest – thus inducing a correspondent variations in the single-tag frequency. On the other hand, the light-jet efficiency is poorly affected. This was expected, as it would be necessary to mistag exactly the additional tagged events, and it would occur in about $10\% \times 5\%$ of events.

The primary vertex search is not a relevant issue here, since the identification of particle emerging from the vertex is performed indirectly. Therefore, the micrometric uncertainties on this point is not considered.

The impact of all relevant uncertainties related to detector, on the four important background processes, are displayed in Tab. 4.1, Tab. 4.2, Tab. 4.3, Tab. 4.4. Not all the listed variables will enter the background uncertainty.

4.1.2 The effect of theoretical systematic bias

Since the present analysis is aimed to estimate the sensitivity reach of a future experiment, it is completely based on simulated data. Therefore, the simulation of physics events occurring in the detector has to resemble real processes as much as possible, and all the parameters involved in the generation of processes have to be properly tuned. On the other hand, many parameters are subjected to fluctuations, due to uncertainty from underlying theories, model dependencies or experimental values input, or may change with luminosity, bunch space or other machine conditions. It is natural to ask how much the present analysis is sensitive to variations in such parameters.

The propagation of theoretical systematic effects on the physics measurements is esti-

Table 4.1: The relative uncertainties produced on a list of variables for the $t\bar{t} \to (Wb)(W\bar{b})$, due to this set of systematic effects: tracker momentum scale (TMS), tracks efficiency and muon efficiency, ECAL energy scale EES, b-tag uncertainty for two different scales, jet energy scale JES for two different scales. In parenthesis, the results with assumptions for the high luminosity (HL) phase are quoted when different from the low luminosity (LL).

	Lepton Eff.	E_T^{miss} Resolution	b-jet Eff.
TMS	1%(0.5%)	-	1%(0.5%)
Tk Eff.	1%	-	1%
μ Eff.	1%	-	1%
EES	0.5%	-	-
b-tag (β_{BTAG}=5%)	-	-	+10%/−15%
b-tag (β_{BTAG}=2%)	-	-	+8%/−10%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=5%)	-	3%	15%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=3%)	-	1%	10%
TOTAL Det. Syst. (LL)	1.8%	3%	+18%/−21%
TOTAL Det. Syst. (HL)	1.6%	1%	+13%/−14%

	1 b-tagg fraction	light-jet Eff.	single jet fraction
TMS	-	0.5%(0)	-
Tk Eff.	-	1%	-
μ Eff.	-	1%	-
EES	-	-	-
b-tag (β_{BTAG}=5%)	5%	0.5%	-
b-tag (β_{BTAG}=2%)	3%	-	-
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=5%)	5%	15%	2%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=3%)	5%	5%	2%
TOTAL Det. Syst. (LL)	7%	15%	2%
TOTAL Det. Syst. (HL)	6%	5%	2%

Table 4.2: The relative uncertainties produced on a list of variables for the $Zb\bar{b} \to (ll)(b\bar{b})$, due to the same set of systematic effects. In parenthesis, the results with assumptions for the high luminosity phase are quoted.

	Lepton Eff.	E_T^{miss} Resolution	b-jet Eff.
TMS	1%(0.5%)	-	1%(0.5%)
Tk Eff.	1%	-	1%
μ Eff.	1%	-	1%
EES	0.5%	-	-
b-tag (β_{BTAG}=5%)	-	-	+13%/−15%
b-tag (β_{BTAG}=2%)	-	-	+7%/−9%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=5%)	-	3%	15%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=3%)	-	-	6%
TOTAL Det. Syst. (LL)	2%	3%	+20%/−21%
TOTAL Det. Syst. (HL)	2%	1%	+9%/−11%

	1 b-tagg fraction	light-jet Eff.	single jet fraction
TMS	-	0.5%(0)	-
Tk Eff.	-	1%	-
μ Eff.	-	1%	-
EES	-	-	-
b-tag (β_{BTAG}=5%)	5%	-	-
b-tag (β_{BTAG}=2%)	2%	-	-
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=5%)	5%	15%	2%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=3%)	1%	5%	1%
TOTAL Det. Syst. (LL)	7%	15%	2%
TOTAL Det. Syst. (HL)	2%	5%	1%

Table 4.3: The relative uncertainties produced on a list of variables for the $t\bar{t} \to l + X$ inclusive production, due to the same set of systematic effects.

	Lepton Eff.	E_T^{miss} Resolution	b-jet Eff.
TMS	1%(0.5%)	-	2%
Tk Eff.	1%	-	1%
μ Eff.	1%	-	1%
EES	0.5%	-	-
b-tag (β_{BTAG}=5%)	-	-	+10%/−15%
b-tag (β_{BTAG}=2%)	-	-	+8%/−11%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=5%)	-	4%	15%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=3%)	-	2%	10%
TOTAL Det. Syst. (LL)	1.8%	4%	+18%/−21%
TOTAL Det. Syst. (HL)	1.6%	2%	+13%/−15%

	1 b-tagg fraction	light-jet Eff.	single jet fraction
TMS	-	1.5%(1%)	-
Tk Eff.	-	1%	-
μ Eff.	-	1%	-
EES	-	-	-
b-tag (β_{BTAG}=5%)	7%	1.5%	0.5%
b-tag (β_{BTAG}=2%)	4%	0.5%	-
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=5%)	5%	15%	1%
JES ($\beta_{JES}(20\,\mathrm{GeV}/c)$=3%)	2%	7%	1%
TOTAL Det. Syst. (LL)	9%	15%	2%
TOTAL Det. Syst. (HL)	5%	7%	1%

Table 4.4: The relative uncertainties produced on a list of variables for the single-top production, due to the same set of systematic effects.

	Lepton Eff.	E_T^{miss} Resolution	b-jet Eff.
TMS	0.5%(0)	-	1%(0.5%)
Tk Eff.	1%	-	1%
μ Eff.	1%	-	1%
EES	0.5%	-	-
b-tag (β_{BTAG}=5%)	-	-	+10%/−20%
b-tag (β_{BTAG}=2%)	-	-	+8%/−15%
JES (β_{JES}(20 GeV/c)=5%)	-	3%	10%
JES (β_{JES}(20 GeV/c)=3%)	-	-	5%
TOTAL Det. Syst. (LL)	2%	3%	+14%/−22%
TOTAL Det. Syst. (HL)	1%	-	+10%/−16%

	1 b-tagg fraction	light-jet Eff.	single jet fraction
TMS	-	-	-
Tk Eff.	-	1%	-
μ Eff.	-	1%	-
EES	-	-	-
b-tag (β_{BTAG}=5%)	5%	-	-
b-tag (β_{BTAG}=2%)	3%	-	-
JES (β_{JES}(20 GeV/c)=5%)	5%	15%	1%
JES (β_{JES}(20 GeV/c)=3%)	3%	5%	1%
TOTAL Det. Syst. (LL)	7%	15%	1%
TOTAL Det. Syst. (HL)	4%	5%	1%

mated below. This estimation firmly assumes that a measure of background from data (that will be described in Sec. 4.2) is affordable. In principle, if all the relevant background is evaluated from a region that is exactly complementary to the signal region, none of the systematic effects in the background have to be taken into account, because uncertainty in simulated events are not an issue anymore. Actually, 'control' regions will be optimized in order to obtain high statistic and low overlapping of different sources, thus they have to be rescaled to the signal region using efficiency factors from simulation, and systematic effects may affect these factors. Additionally, theoretical systematic effects may lightly alter the shape of background distribution, leading to fluctuations in the fit performed to count events. Since in most of cases the phase space from where signal is extracted is much smaller than the one from which background is estimated, these effects are generally negligible.

Hereafter, the impact of theoretical uncertainties will be evaluated only for signal selection efficiency and surviving background events from $t\bar{t}$ (both di-leptonic and inclusive channels) and single-top in the t-channel. Other minor sources as $Zb\bar{b}$ are such that they can be directly estimated in the same signal window (properly enhanced in background content), thus a parametrization is not needed. As shown in the previous chapter, last cuts in that window strongly suppress these contributions, and theoretical fluctuations are not expected to play a role.

Most of strategies to assess the impact of theoretical uncertainties are based on the analysis in Ref. [140], that CMS collaboration has managed to contain valuable guidelines.

When a new generation and simulation step with parameters different from the analysis sample has required, the fast simulation package FAMOS (outlined in Sec. 2.3.5) has been adopted. It allows a quick generation of events and all the parameters can be easily set by the user. In some cases, statistic is still limited to few thousands of events: on the other hand, systematic effects having a sizable impact on the analysis should emerge on top on fluctuations induced by this statistic.

Uncertainties in cross section

In general, the knowledge of the total and differential cross sections is limited, because it is influenced by the order at which calculations are performed and some of the parameters outlined below. Since the theoretical uncertainty on the cross-sections values, in many cases, exceeds 10-20%, some techniques to estimate background from data are welcome. In that way, absolute cross-section fluctuations and luminosity instabilities are not an issue anymore. The error on the cross section for the $t\bar{t}$ production will become smaller and smaller as the process will be directly observed at LHC.

On the other hand, distortions in differential distributions may be induced from variations in several parameters. Most relevant ones are sketched in the following.

QCD scale uncertainty

In the generation of top production process the hard scale (Q^2) has been tuned to the $m_t^2 + p_T(t)^2$ value. This parameter set the value of α_S and the running of PDF, hence the behaviour of the cross sections may be affected. The dependence of the observables on this scale is unphysical (since it is only a cut-off for calculations) but, when events are generated at lowest order as in PYTHIA, cross sections may be significantly affected from the choice. On the other hand, only an effect around percent level on $p_T(t)$ distribution is expected (see for instance Ref. [22]). This variable has been never used in the analysis: in fact, the most important background have a top in the final state, so any difference in top momentum between signal and background is supposed to exist. For that reason, scale uncertainty are expected to produce no sizable effects in the full analysis and it is ruled out from the present study.

Initial and final state radiation

In principle, all the particles involved in the event and subjected to strong interaction may emit a gluon radiation, in the initial or in the final state (Initial/Final State Radiation, ISR/FSR). These effects give rise to extra jets and cause the kinematics of the reconstructed objects to be different than expected.

The analysis here is strongly based on b- and light-jet reconstruction, thus an impact should in principle be sizable. While the high cut in the transverse momentum of light-jet rules out a significant amount of gluon radiation (that accumulates a low energy), some effects might be expected on the b-jet multiplicities (as these jets have a lower p_T cut) and on an additional smearing of η and ϕ direction of jets. Contribution on mis-identification of leptons and photons (that are selected after strong isolation requirements) should be negligible.

In Ref. [22] it has been show that, once the proper matrix elements for gluon emission from b quark are introduced in the generator, final state radiation does not alter the jet kinematics and multiplicity. On the other hand, ISR still could play a role.

Following Ref. [140], the switches regulating the amount and the strength of the radiation of the showering has been varied. Namely, the virtuality scale Λ_{QCD} up to which ISR stops and from which FSR starts is varied between 0.150 and 0.280 GeV, for space-like and time-like showers together, and a set of five different signal samples are generated. The four variables leptons efficiency, b-tagging and light-jet identification efficiency, fraction of single-tagged jet are then compared and results are presented in Tab. 4.5. They are all found consistent within few percent, with the exception of b-jet efficiency. When statistical fluctuations are disentangled, a fluctuation in the b-jet number as large as $\sim 20\%$ is still there – even if this happens only for the largest Λ_{QCD} value. On the other hand, similar motivations rise to not include this fluctuation in the analysis: in fact, since it can be ascribed only to ISR, it is expected to occur equally in all kinds of $t\bar{t}$ production. In both analyzed channels this

Table 4.5: Comparison of four benchmark observables in the signal selections, when five samples with different ISR/FSR parameters are generated.

Λ_{QCD}	Lepton Eff.	b-jet Eff.	1 b-tagg fraction	light-jet Eff.
0.150	0.124 ± 0.007	$(6.2 \pm 1.2) \cdot 10^{-3}$	0.75 ± 0.29	$(7.7 \pm 1.3) \cdot 10^{-3}$
0.183	0.126 ± 0.008	$(6.7 \pm 1.2) \cdot 10^{-3}$	0.87 ± 0.32	$(8.4 \pm 1.4) \cdot 10^{-3}$
0.216	0.125 ± 0.008	$(5.8 \pm 1.1) \cdot 10^{-3}$	0.77 ± 0.32	$(9.2 \pm 1.5) \cdot 10^{-3}$
0.250	0.124 ± 0.008	$(4.2 \pm 1.0) \cdot 10^{-3}$	0.79 ± 0.27	$(7.8 \pm 1.3) \cdot 10^{-3}$
0.283	0.129 ± 0.007	$(8.5 \pm 1.5) \cdot 10^{-3}$	0.57 ± 0.20	$(5.8 \pm 1.2) \cdot 10^{-3}$

constitutes the main background, so the same bias is expected: as in the previous case, it will be rejected in the data samples when background will be subtracted from the signal region.

Top input mass

The mass of top quark is a parameter that is needed in many theoretical calculations (even at LO), as those for production cross section, mass width, polarization and others. Among the input parameters, it is by far the one known with less accuracy: not only its experimental error has slided down 10% only recently (as mentioned in Sec. 1.1.1) but the declared central value is about 3-5 GeV lower than just one year earlier [23] and than the value $m_t = 175\ \mathrm{GeV}/c^2$ used in generators like TOPREX.

Though the $t\bar{t}$ cross section is known to strongly depend of the mass (with a rule of thumb $\delta\sigma/\sigma \sim 5\,\delta m_t/m_t$) no measurable effects are foreseen for kinematic variables [22]. Also in the single-top case, no significant modifications of the production kinematics are expected, since the differential cross section for the t-channel is modulate by the momentum $d\sigma/q^2 \sim 1/\left(q^2 - M_W^2\right)$.

Moreover, it is important to remark that signal regions in the top invariant mass distributions are large enough to accommodate all the reasonable top mass adjustments. As a result, both the distributions of top observables and the analysis cuts can be safely consider unaffected from m_t uncertainty, and this issue is not considered anymore.

Light and b quark fragmentation

After being produced from the top quark, W boson, proton partons or ISR/FSR effects, the collection of quarks and gluons hadronizes into mesons and baryons. The fraction of the original b quark momentum which will appear as visible energy in the reconstruction cone of the corresponding b-jet depends on the 'fragmentation' function of the b quark. As this analysis involves a signal and a background with c- and b- jets, the description of the heavy quarks fragmentation may be relevant.

In generator codes, the process is described by a set of 'fragmentation' parameters that

Table 4.6: Efficiencies for three benchmark variables in the FCNC signal selection. Results reflect the effect of changing the a and b parameter of the Lund fragmentation model.

Observed variable	$a + \Delta a$, $b + \Delta b$	a, b	$a - \Delta a$, $b - \Delta b$
b-jet Eff.	0.382 ± 0.010	0.322 ± 0.010	0.366 ± 0.010
1 b-tagg fraction	0.96 ± 0.16	0.90 ± 0.17	0.89 ± 0.16
light-jet Eff.	0.609 ± 0.017	0.603 ± 0.017	0.521 ± 0.015

cannot be obtained in a perturbative way. The non-perturbative description of fragmentation is realized via models, which need to be adjusted to agree with experimental data. If fragmentation is assumed to depend only on the factorization scale (*i.e.*, jets fragment in the same way at hadron and lepton machines), once the parton shower cut-off scale is fixed, the LEP/SLD tunings (or the Tevatron ones) can be used as they were for the LHC.

PYTHIA performs fragmentation by splitting gluons into quark-anti-quark pairs, then turns the resulting set of colour singlet into hadrons via a string model. For the b-jet, the string (or Lund) fragmentation model [132] or the Peterson function [133] are commonly used.

Following suggestions in Ref. [140], the two (correlated) parameters of Lund model are varied by their uncertainty ($a = 0.11 \pm 0.04$ GeV^{-1}, $b = 0.55 \pm 0.04$ GeV^{-1}) along with the width of transverse momentum distributions for primary hadrons (0.36 ± 0.03 GeV). These values hold when $\Lambda_{QCD} = 0.250$ GeV and Q^2_{min} (the invariant mass below which partons are not assumed to radiate) are equal to 1.90 GeV (from a fit by OPAL).

Even if the comparison is limited to three values only, an additional effect could be inferred for both the b and the light-jets. If the efficiency values are assumed to be Gaussian distributed and consistent within 2σ (95.5%) with no systematics, a shift has to be added to be consistent with the result, that corresponds to about 9% for the b-jets and 11% for the light ones. The same effect in tagging efficiency is to be expected for all the top background and for $Zb\bar{b}$ as well, since the b-jets that enter the signal region features the same kinematic. Differently from the previous cases, this effect influences the final state, thus it cannot be subtracted anymore from the signal and has to be considered as an additional systematic uncertainty.

Figure 4.1 displays the fluctuation on the $p_T(b\text{-jet})$ produced from this theoretical uncertainty.

Parton density function

The suggested way to evaluate theoretical uncertainties due to a certain proton PDFs is to vary the errors on the parameters of the PDF fit itself. With the Les Houches accord [141] PDF (LHAPDF) errors should be easily propagated via re-weighting to the final observables.

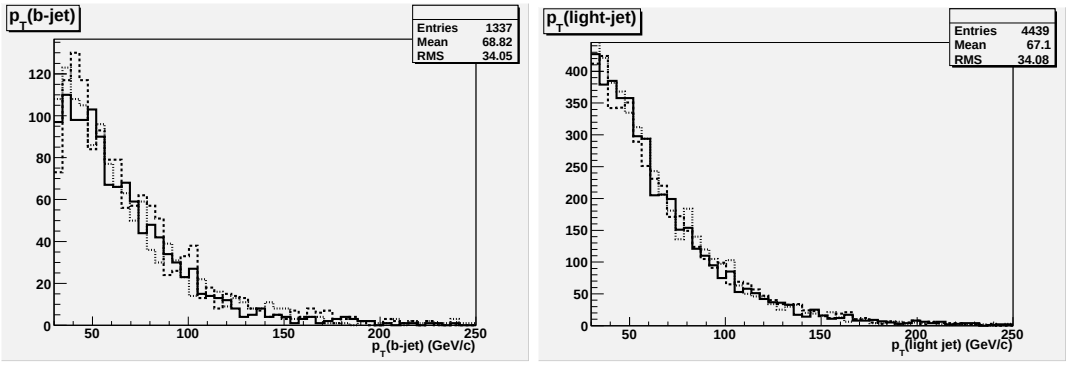

Figure 4.1: A comparison between three signal samples generated with different b and light quark fragmentation parameters, namely a, b (solid line), $a + \Delta a$, $b + \Delta b$ (dashed line), $a - \Delta a$, $b - \Delta b$ (dotted line). Distribution of transverse momenta for light (right) and b-jets (left) are represented. Units on the vertical scale are arbitrary and statistic is the same.

For analyses which are known to be particularly sensitive to PDFs, like cross-section measurements, it would be desirable to compare two different sets of PDFs (typically CTEQ vs. MRST) taking then the maximum variation as an extra error. This is important since, even considering the error boundaries, different set of PDFs may not overlap in some region of the phase space.

With the peculiar configuration of the signal and background in the present analysis, the estimation of this kind of systematic effect may be considered unessential. The statement is motivated to one side by the fact that the $t\bar{t}$ background is the only really important background source in both channels, to the other side that mostly of the background fluctuations will be measured directly from data. On this topic, the following considerations can be set up:

- the probability density functions involve the kinematic distributions of the partons in the proton, so they are only able to affect the $t\bar{t}$ production. The signal and the top pair background, that are different by their final states, do not differ at tree level for the effect of a fluctuation in a PDF;

- if a PDF uncertainty results in a modification of $t\bar{t}$ cross-section value, the case discussed above presents so it is not an issue. A cross-section fluctuation will be directly measured from data, and it will be not considered a fluctuation anymore;

- if, for some particular choice of the PDF parameters, the multiplicity of the final state is affected (number of leptons, number of jets or b-jets) and the modified final state survives up to the last cut, a process affecting both the signal and the major background is at work. If the background estimate is properly subtracted when searching from the signal, all the additional particles that has developed are completely removed;

- if a PDF uncertainty modifies the shape of the background distribution, the effect on

the extrapolation from control to signal region is negligible, because the signal is often sampled on a very small range with respect to the interval on where background is estimated. In the present analysis, this will be true for the $t \to qZ$ study.

The conclusion is that, if the background will be estimated through the techniques proposed in the following, all the effects related to the imperfect knowledge of PDF parameters cancel, or do not play a significant role. Hence this source of systematic errors will be not included.

Considerations about the underlying event

The underlying event (UE), composed of the protons remnants and the effects of multiple interactions, also contributes to extra jets in the event[1].

The composite nature of collisions allows for several partons from each of the incoming hadrons to undergo scatterings. Such multiple parton-parton interactions are instrumental in building up the activity in the underlying event, in everything from charged multiplicity distributions and long-range correlations to mini–jets and jet pedestals.

To assess the impact of UE on jet selection in this analysis, strategies have to be followed to reproduce the "pedestal effect" found at Tevatron, *i.e.* a raise of the number of charged particle tracks per unit of pseudorapidity when the underlying event is at work. This can be done at generator level by varying the impact parameters for the parton-parton interactions and, with PYTHIA, Tevatron effect can be extrapolated at 14 TeV.

In the present analysis, such effects are not dissimilar from those related to PDF uncertainties, thus one can get rid of them with the same motivations.

No reasons can be devised such as the underlying event hidden in the $t\bar{t} \to (Vb)(W\bar{b})$ process is different from the one in the $t\bar{t} \to (Wb)(W\bar{b})$ process. Therefore, this extra activity is expected to be detected when estimating background from data; if a portion of it falls in the signal region, it will be subtracted from the signal, that is affected by the same activity, and cancels out. If that strategy for background counting is adopted, the analysis results insensitive also to this uncertainty source.

Minimum bias and pile-up effects

As mentioned more than once, in a hadron collision the event in which one is interested is usually accompanied by multiple inelastic collisions, that occur in the same bunch crossing. These events are of concern here, because they may give rise to extra jets originating from other primary vertexes. In addition, the flow of additional soft charged particles which spiralize in the high magnetic fields of CMS may have an effect on the tracker detectors,

[1] Since the impact of the initial state radiation has already been accounted for, here the processes that "stays under" the main event does not contain the ISR.

Table 4.7: Some relevant efficiencies for the $t \to Zq$ signal selection are compared for two scenarios with different pile-up. Poissonian means of the number of minimum bias (MB) are set to 3.5 and 7 for each hard scattering event.

Observed variable	3.5 MB evs.	7 MB evs.
Z^0 Reconstruct Eff.	0.184 ± 0.006	0.182 ± 0.006
b-jet rec. Eff.	0.356 ± 0.009	0.340 ± 0.009
1 b-tagg fraction	0.87 ± 0.42	0.91 ± 0.36
anti-tagged jet rec. Eff.	0.080 ± 0.004	0.072 ± 0.004

where multiple hits in the same layer can be generated by the same track. Consequently, b-tagging properties can be altered and the lepton efficiency may worsen.

The number of minimum bias interactions generated in a single beam crossing is a Poissonian distribution that depends on the instantaneous luminosity, which varies of about a factor 2 during a LHC fill. The approach followed here has been to compare a sample of signal events analyzed at $10\,\mathrm{fb}^{-1}$, where the mean number of such events are 3.5, with a sample having twice the amount of minimum bias. This permits to estimate a systematic error that conservatively takes into account the largest possible pile-up fluctuations during a run.

Results are quantified in Tab. 4.7. The impact of a possible reduction of lepton efficiency is measured via the Z invariant mass, by comparing the number of reconstructed objects. The invariant mass is reconstructed in a way quite similar to the analysis step, with exception of a lightly softer isolation ($Isol(e/\mu) < 0.1$) and no cuts in the mass window. The number and multiplicity of tagged jet, with $p_T > 30\,\mathrm{GeV}/c$ and discriminator greater than 2 are compared. Only the fractions of leading jets harder than $30\,\mathrm{GeV}/c$ and with a negative discriminator are compared.

Although statistic is rather limited, reasonable motivations exist to state that systematic uncertainties in the lepton and b-tagging efficiency are well below 1% and 2%, respectively. While the fraction of b-tagged objects is not a handle to discover possible effects, the 2σ distance in the efficiency for the anti-tagging could trace something. In any case, a systematic bias event in this selection should not be greater than $\sim 10\%$.

This exercise can be intended as preliminary to the extrapolation to much more severe pile-up scenarios. In the next section, the challenges posed by the high luminosity phase of LHC and by Super LHC will be partially simulated, and the same observables will be compared.

Charge Asymmetries

All the selection procedures detailed in the previous chapter has been applied to a signal that was symmetric to charge conjugation, *i.e.* with equal admixtures of $t\bar{t} \to (bW^+)(\bar{q}Z/\gamma)$

and $t\bar{t} \to (qZ/\gamma)(\bar{b}W^-)$. There are no physics reasons to favour one state with respect to charge conjugate one, since to present knowledge the production and decay rates of top are irrespective to its charge and, consequently, kinematic of the final state particles respects C-parity. At the lowest order, the amount of energy loss and multiple scattering does not depend from the charge sign, and the detectors response should be symmetric for two different sign states. Finer effects as different energy loss of kaons state and possible muon detector asymmetries are pretty lower than the sensitivity of this search. As a result, each asymmetry noticed in the results should be ascribed to bugs in the generator code or in the simulation framework.

To test the charge symmetry of the analysis, the signal sample is partitioned in two conjugate charge samples and the selection is repeated for both.

For the $t/\bar{t} \to q/\bar{q}Z$ channel, the set of selection considered is:

- the presence of a $Z \to l^+l^-$, where $l = e, \mu$;
- the presence of a $Z \to l^+l^-$ plus a third lepton and missing energy. Testing these two efficiencies may reveal possible biases in the lepton detection;
- the combined efficiency for a W, a Z and a tagged b-jet, that takes into account the b-tagging efficiency;
- the presence of a top with standard decay and a hard light-jet, that is sensitive to the jet reconstruction algorithm.

A similar set of cuts is devised for the $t/\bar{t} \to q/\bar{q}\gamma$ channel, namely:

- the reconstruction efficiency for $W^\pm \to l^\pm\nu$, where $l = e, \mu$;
- the presence an additional b-jet, to examine the tagging charge symmetry even in this case;
- the presence of a top with standard decay, as a further check of charge invariance reconstruction .

As only the response to charged objects is of concern here, the photon selection efficiencies are not compared. The results are represented in Tab. 4.8 and Tab. 4.9 with their statistical errors.

As expected, all fluctuations between the two samples are contained in one sigma, thus no evidence can be drained of visible charge asymmetries. Since in all the background sources the charged objects to be reconstructed are basically the same (leptons, jets, b-jets), the conclusion is that this effect, if present, is not at all relevant in the analysis.

Table 4.8: Partial efficiencies for a subsample of the signal $t \to Zq$, for the two charge conjugate states. Errors are statistical only.

Observed variable	$t\bar{t} \to (bW^+)(\bar{q}Z)$	$t\bar{t} \to (qZ)(\bar{b}W^-)$
Z rec. Efficiency	$(42.8 \pm 1.0)\%$	$(42.8 \pm 1.2)\%$
Z+W rec. Efficiency	$(20.4 \pm 0.7)\%$	$(20.2 \pm 0.8)\%$
Z+W+b-jet rec. Efficiency	$(6.2 \pm 0.4)\%$	$(5.8 \pm 0.4)\%$
t_{SM} + q-jet rec. Efficiency	$(2.1 \pm 0.2)\%$	$(1.8 \pm 0.2)\%$

Table 4.9: Partial efficiencies for a subsample of the signal $t \to \gamma q$, for the two charge conjugate states. Errors are statistical only.

Observed variable	$t\bar{t} \to (bW^+)(\bar{q}\gamma)$	$t\bar{t} \to (q\gamma)(\bar{b}W^-)$
W rec. Efficiency	$(34.1 \pm 0.9)\%$	$(33.85 \pm 1.3)\%$
Z+W+b-jet rec. Efficiency	$(8.6 \pm 0.5)\%$	$(9.3 \pm 0.7)\%$
t_{SM} rec. Efficiency	$(6.9 \pm 0.4)\%$	$(7.1 \pm 0.6)\%$

4.1.3 Summary of systematic errors contributions

At this level, enough results have been collected that allow to summarize the cumulative impact of all the systematic effects. In the following tables, these effects are reported as relative fluctuations only on the variables that will demonstrate to be relevant when scaling is done from the control region (where background is counted) to the signal region, where signal efficiency is measured.

The estimation is further divided in two parts:

- one for a 'low luminosity' scenario, approximately defined as the phase from 10 to $50\,\mathrm{fb}^{-1}$ integrated luminosity. Here all references values for detector errors are assumed: namely, tracker and muon chamber misalignments have the impact detailed above, b-tagging is not known better than 5% and jet energy scale precision ranges from 5 to 3%;

- one for a 'high luminosity' scenario, that is intended to extend from $50\,\mathrm{fb}^{-1}$ to the whole accelerator life. At that time, enough statistic from data should be drained, that ensures a negligible impact both from misalignments, muon scale and ECAL scale. Jet energy scale fluctuations are assumed to go from 3 to 1.5% and b-tagging to be understood ad the 2% level. Errors on the limited knowledge of the B-field, though small, are supposed to be still there.

This division clearly do not correspond to an abrupt change between two scenarios. On the contrary, it is only proposed to reproduce a situation for which predictions are well

Table 4.10: The cumulative impact of systematic effects, both from detector and from theory, on some variables that influence the amount of background. Here the parameters for systematic estimation are those expected from the very first tens of integrated inverse femtobarns.

Background source	Lepton Eff.	E_T^{miss} Resolution	b-jet Eff.
$t\bar{t} \to (l\nu b)(l\nu b)$	1.8%	3%	+20%/ − 23%
$Zb\bar{b} \to (ll)(b\bar{b})$	2%	3%	+22%/ − 23%
$t\bar{t} \to l + X$	1.8%	4%	+20%/ − 23%
$qt \to (l\nu b) + X$	2%	3%	+17%/ − 24%
Background source	1 b-tagg fraction	light-jet Eff.	single jet fraction
$t\bar{t} \to (l\nu b)(l\nu b)$	7%	19%	2%
$Zb\bar{b} \to (ll)(b\bar{b})$	7%	19%	2%
$t\bar{t} \to l + X$	9%	19%	2%
$qt \to (l\nu b) + X$	7%	19%	1%

motivated in a conservative but realistic way, together with a situation where most of target results should be achieved. No account is given, for instance, to a possible re-commissioning phase related to the replacement of many detectors for SLHC.

Systematic effects related to the limited theoretical knowledge are instead considered constant in all the luminosity phases.

For what concerns theoretical uncertainties, as it was explained only the effect of b and light fragmentation should have some impact. It is important to note that the estimate has been performed with less-refined tagging and jet identification algorithms, thus some improvements in robustness may be still possible. An increase of statistic for the analysis of the effect should give rise to a situation where the error from this source is reduced. On the other hand, its order of magnitude is assumed to be fairly correct.

4.1.4 Outlook to LHC at high $\mathcal{L}$ and SLHC

Given the large top quark cross section, most of the top physics programme is intended be completed during the first few years of LHC operation. The measures of the $t\bar{t}$ cross section and top mass, for instance, are supposed to steadily approach to the theoretical uncertainties, reaching an uncertainty beyond which more data offer no obvious improvement.

The analysis contained in this work, on other hand, addresses no common processes so it is still limited by the statistic. An increase of data size delivered by the machine will naturally lead an improvement of statistical precision and, if a FCNC process would not be discovered, a better and better upper limit will come with the data. In this respect, it make sense to briefly address the detection of these anomalous decay at LHC with high luminosity (10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$) and even at the designed LHC improvement, called Super LHC (SLHC). Some

Table 4.11: The same summary of systematic errors is performed as in Tab. 4.10, now with some reduced detector uncertainties as foreseen for a more advanced age. Theoretical errors have demonstrated to be relevant only in the jet efficiency, and are assumed constant all over the LHC phases.

Background source	Lepton Eff.	E_T^{miss} Resolution	b-jet Eff.
$t\bar{t} \to (l\nu b)(l\nu b)$	1.6%	1%	$+16\%/-17\%$
$Zb\bar{b} \to (ll)(b\bar{b})$	2%	1%	$+13\%/-14\%$
$t\bar{t} \to l + X$	1.6%	2%	$+16\%/-17\%$
$qt \to (l\nu b) + X$	1%	-	$+13\%/-18\%$
Background source	1 b-tagg fraction	light-jet Eff.	single jet fraction
$t\bar{t} \to (l\nu b)(l\nu b)$	6%	12%	2%
$Zb\bar{b} \to (ll)(b\bar{b})$	2%	12%	1%
$t\bar{t} \to l + X$	5%	11%	1%
$qt \to (l\nu b) + X$	4%	12%	1%

insights on this machine are sketched in the Appendix, Sec. A.3.

Estimates of benchmark efficiencies

An attempt to check the reconstruction of some high-level objects, in two different luminosity scenarios, is outlined below. Three subsamples have been generated with different pile-up: the low luminosity with 3.5 minimum bias events, the high luminosity with 17 minimum bias events, and a possible SLHC scenario where the pile-up is tuned on 230 minimum bias events (as indicated for instance in Ref. [142]). For the three samples, five different observables are compared:

- the efficiency for the reconstruction of a Z boson, from a pair of muons and electrons. In one shot, it gives an idea of both the lepton identification efficiency (roughly corresponding to the square root of the Z efficiency) and the mass resolution of the analysis. Both electrons and muons are pre-selected with the same criteria exploited in the analysis step; isolation variable is considered in the same cone, but with $Isol(e) = Isol(\mu) < 0.1$;

- the peak value of the reconstructed $M(ll)$ invariant mass, simply taken as the centroid of the histogram. A possible bias in this value could shed light on a modification in the lepton energy scale, when upgrading to higher luminosities;

- the reconstruction efficiency for the b-tagged jets, that is instrumental in ensure a good signal over background ratio. Tagging algorithm is tuned on the same discriminator cut than in the analysis, while selection includes jet down to $30\,\mathrm{GeV}/c$ – since here maximizing the efficiency is not an issue for a comparison;

Table 4.12: A comparison of some reconstructed variables for three LHC phases, characterized by different intensity of pile-up – in other words three different values for the amount of minimum bias (MB) events. They are averagely 3.5 for LHC at low luminosity, 17 for LHC at low luminosity and about 230 for SLHC. Errors are statistical only and the meaning of efficiencies is detailed in the text.

Observed variable	LHC Low Lumi (3.5 MB evs.)	LHC Low Lumi (17 MB evs.)	SLHC (230 MB evs.)
Z^0 Reconstruct Eff.	0.184 ± 0.006	0.179 ± 0.006	0.172 ± 0.006
Z^0 Mean Value	90.53 ± 0.15	89.78 ± 0.15	87.64 ± 0.15
b-jet rec. Eff.	0.356 ± 0.009	0.336 ± 0.009	0.351 ± 0.009
1 b-tagg fraction	0.87 ± 0.14	0.90 ± 0.15	0.83 ± 0.13
anti-tagged jet Eff.	0.080 ± 0.004	0.079 ± 0.004	0.101 ± 0.005

- the fraction of a single-tagged jet per events, with respect to all the tagged jets. This number could be sensitive to the efficiency for the double b-tagging (that in principle differs from the tagging efficiency) and directly enters the estimation of the $Zb\bar{b}$ background for the $t \to qZ$ signal. This topic will be detailed in next section;

- the efficiency for the detection of light-jets, that can be influenced from fluctuations in the jet energy scale. In some of the cases presented here, light-jet efficiency affects the background estimation because it discriminates the signal region from the control region in which background is counted. Instead of setting up the jet cleaning procedure adopted in the analysis, here non-b-jets are simply taken as the objects tagged with a negative discriminator by the reconstruction framework. Though efficiency is in principle different from the analysis, and u and c components are not so well controlled, a comparable effect from detector systematic bias is expected.

Results of the exercise are presented in Fig. 4.2 for the comparison of the two LHC luminosity phases, Fig. 4.3 for the comparison of LHC high luminosity phase with SLHC, and in Tab. 4.12 all together.

Although the minimum bias increases by two orders of magnitude, no major effects seem to emerge. Most of these "benchmark" variables present fluctuations that nicely accommodate within the statistical errors, thus possible distortions would affect only below the percent level. Namely, even with the huge pile-up increasing, momentum and ECAL resolution as well as lepton energy scale should have at most a very little distortion. The number of b-tagged jets and their multiplicity distribution, even if affected by a larger statistical error, are well consistent within 2σ. The largest effects in shifting from LHC to SLHC are an increase about 18% in the total number of the anti-tagged jets: if the anti-tagging efficiency for 17 MB is supposed to be equal with that at 230 MB within 2σ, some 15% systematic bias should be added. As Fig. 4.3 testifies, the discrepancy strongly depends from the chosen p_T cut in

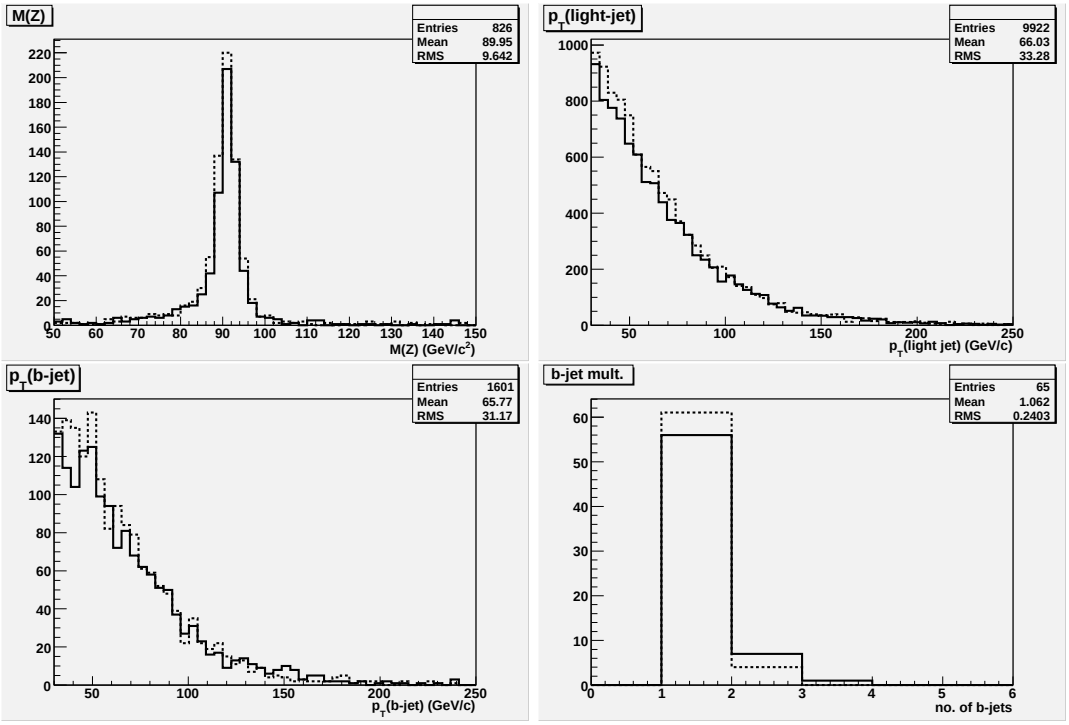

Figure 4.2: The reconstruction of the invariant Z mass (left upper), the p_T distribution for the anti-tagged jets (right upper), the p_T distribution for the b-tagged jets (left lower) and their multiplicity distribution (right lower), for a sample of signal event with low luminosity pile-up (3.5 minimum bias event, solid line) and high luminosity pile-up (17 minimum bias event, dashed line). The two samples have the same number of events.

the selection. It can be shown that with p_T(light-jet) $>80\,\mathrm{GeV}/c$, discrepancy falls below the 5% level, so being completely drowned by the JES systematic error.

On the other hand, the analysis results cannot be directly extrapolated up to SLHC luminosity, since in that environment the trigger menu can be fairly different. Studies on possible trigger paths for these high pile-up scenarios are ongoing and no simulations can be still exploited. While in such scenarios quite harder objects will be selected already at the trigger stage – hence getting rid of the light-jet excess – the efficiencies for signal and background could be significantly reduced. At this level, no predictions can be inferred about the sensitivities to the addressed channels in this context: as a consequence, extrapolations will be performed only up to $100\,\mathrm{fb}^{-1}$, where no major modifications for the trigger parameters are foreseen.

4.2 Estimation of background from data

In the approach followed in the previous chapter, the amount of surviving background in both channels has been evaluated by counting the number of event filling the last selection. Though it is the simpler solution, the background number results to be affected by several

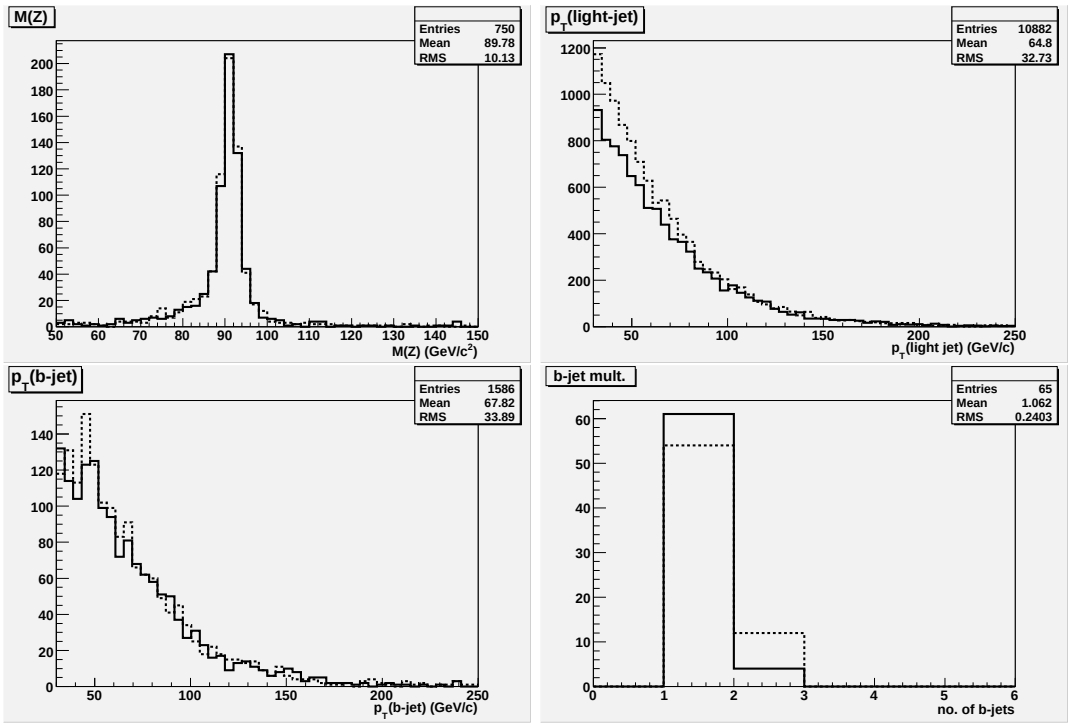

Figure 4.3: The reconstruction of the invariant Z mass (left upper), the p_T distribution for the anti-tagged jets (right upper), the p_T distribution for the b-tagged jets (left lower) and their multiplicity distribution (right lower), for a sample of signal event with the high luminosity pile-up (17 minimum bias event, solid line) and with the pile-up expected for SLHC (230 minimum bias event, dashed line). The two samples have the same number of events.

uncertainties, as:

- high statistical error, due to the small number of background events;
- experimental uncertainties in the luminosity, that is used to obtain the rate from the cross section;
- theoretical error in the cross sections of all processes, affecting their magnitude and their shapes;
- the result of all experimental systematic uncertainties related to each selection cut.

The uncertainty on the background propagates on the significance estimates, thus producing significant oscillations in the final sensitivity limit.

The commonly used solution is to identify a region ideally signal-free and enriched in only one type of background process. This region should be suitable to be searched for in future data samples with the lowest systematic error. The amount of the background evaluated in this region can be subsequently extrapolated to the signal region, using rescaling with efficiency factors as predicted by Monte Carlo. The advantages of this procedure are striking:

- statistic errors come from a highly populated region in the phase space, so they are much reduced;
- the control region is designed to be extracted from data, so it is protected against uncertainties in luminosity and absolute cross sections;
- uncertainties in distributions shapes are usually much smaller than theoretical errors on normalization;
- detector and theoretical systematic effects play a role only in ratios of selection efficiencies, thus their uncertainties can be effectively reduced.

4.2.1 Guidelines for the background estimation

The general expression for the number of events of a background process B_j is determined by counting their number in a control region and extrapolating to the signal region:

$$N_{B_j}^{Sign} = N_{B_j}^{Contr} \epsilon_j^{C \to S} \quad , \tag{4.1}$$

where $N_{B_j}^{Sign}$ ($N_{B_j}^{Contr}$) is the number of event in signal (control) region and $\epsilon_j^{C \to S}$ the reduction factor from one region to the other, estimated by simulated events in the analysis.

The error on $N_{B_j}^{Contr}$ is the combination of statistical error $\sqrt{N_{B_j}^{Contr}}$ (the counting distribution is assumed to be Poissonian) and systematic uncertainties, whose relative effect

is δ_{SYST}. Therefore the total error on $N_{B_j}^{Sign}$, when exploiting this 'extraction from data method' is:

$$\Delta N_{B_j}^{Sign} = \sqrt{\left(\Delta_{SYST}\epsilon_j^{C\to S} N_{B_j}^{Contr}\right)^2 + \left(\epsilon_j^{C\to S}\Delta N_{B_j}^{Contr}\right)^2} \quad ,$$

where the error $\Delta N_{B_j}^{Contr}$ is:

$$\Delta N_{B_j}^{Contr} = \sqrt{\Delta N_{B_j}^{Contr} + \delta_{SYST}^2 N_{B_j}^{Contr\,2}} \quad ,$$

In practical cases some more complications arise, because to find a region where all background sources but one are negligible often is not possible. When that happens, 'contaminating' events have to be counted (from simulated events or from data) and subtracted from the sample. Absolute uncertainty on the subtracted number of events has to be added to $\Delta N_{B_j}^{Contr}$.

The control regions for each background process are chosen in order to have a low contamination from other processes, to minimize the total error $\Delta N_{B_j}^{Contr}$, to have as much rich as possible statistic and to be addressable in future data with a simple choice of datasets. These general guidelines are now applied to both the analyzed processes.

4.2.2 Background estimation for the $t \to Zq$ channel

From the analysis detailed in Chapter 3, $t\bar{t}$ and $Zb\bar{b}$ have resulted as the only interesting background sources. It is useful to briefly remind the main features of these two backgrounds with respect to the signal:

- the signal has a narrow Z^0 peak in the di-lepton invariant mass distribution, only one b-tagged jet and a significant missing energy;
- the $Z + b\bar{b}$ production has the Z^0 peak in the di-lepton distribution, two b-tagged jets and an amount of missing energy similar to the signal;
- the top pair production following di-leptonic SM decay channel has no peaks in the di-lepton distribution[2], has two b-tagged jets and the missing energy distribution has a rather larger tail than the other processes.

These two background sources are estimated separately.

Estimation of the $t\bar{t} \to 2l$ background

It is straightforward to guess that the best measure of this background may be obtained from the side-bands in the e^+e^- and $\mu^+\mu^-$ invariant mass distribution, as they share only a very

[2]The 'bump' originated by the request for the lepton pair closest to the Z nominal mass is very small and has no relevance here.

little number of events with the signal. Since the rate of $t\bar{t}$ is high, these region can be further cleaned from the other components, but a compromise has to be reached between a too mild cleaning (that leaves there the other background sources) and a too severe one (that reduces too much the statistic yield). The following considerations hold:

- asking for at least one b-jet effectively suppresses most of QCD and boson+jet background, thus it is a powerful tool to increase the purity of the control region in $Zb\bar{b}$ and $t\bar{t}$;
- to reduce the $Z + b\bar{b}$ contribution in the control region from $t\bar{t}$, a large lower limit for missing transverse energy is proposed. This turns out to be very powerful in suppressing other sources not included in the analysis, as inclusive $t\bar{t} \to 2l$, single-top, $Z + b + j$ production. The optimal threshold for E_T^{miss} is discussed below;
- the presence of a third hard lepton, with the same kinematical cuts of signal selection, would be instrumental in rejecting most of processes leading to a $b\bar{b}+l^+l^-$ in their final state. On the other hand, this would lead to an unacceptable reduction in the statistic for $t\bar{t}$ (at $10\,\mathrm{fb}^{-1}$, around 20 surviving events if the large missing energy requirement is added): therefore, no requests for a further lepton has introduced in this control region;
- as it was demonstrated, the ask for more than one b-jets has a quite modest effect on the discrimination of the SM from the FCNC signal, because the b-jet selection procedure leads to a tagging multiplicity for the $t\bar{t}$ that is not very different than for signal. Moreover, the previous section has shown that many systematic effects, both theoretical and detector-related, may alter the number of b-jets and the ratio between the number of events with one jet and with more. Consequently, the control region here is chosen with the same b-tagging requirement as the signal region, *i.e.* to have one b-jet only. The great advantage is to avoid suffering from uncertainties in b-tagging, at the low price of a small statistic reduction;
- still in order to avoid systematic effects related to additional cuts, less-efficient selections on $M_T(W)$ or $M_T(bW)$ are not introduced.

In all, the contaminating processes different from $Zb\bar{b}$ and $t\bar{t}$ are considered well suppressed with the two first cuts.

The distribution of the selected E_T^{miss}, obtained in the same way as explained for the signal, is displayed in Fig. 4.4 for the two most relevant backgrounds and one of the possible minor sources. Here they are rescaled to the respective cross sections. As already observed, missing energy due to semi-leptonic b quark decays in the $Zb\bar{b}$ and Z+jets, (coming from semileptonic decays of heavy flavours in jets, or decays of τ from Z) is shifted to lower values with respect to W semi-leptonic decays from top, so optimizing the lower E_T^{miss} cut is easy. The better $t\bar{t}$ purity is found when E_T^{miss} >85 GeV and is quite robust around this value, since to deteriorate this purity by 4% a E_T^{miss} resolution worse that 10% would be necessary.

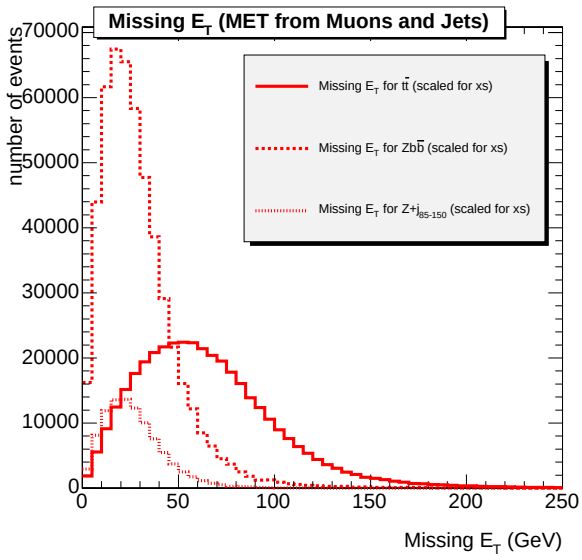

Figure 4.4: The distribution of missing transverse energy for three background sources in the $t\bar{t} \to (Zq)(W\bar{b})$ channel, that can be very useful to disentangle the $t\bar{t}$ from the $Zb\bar{b}$. Best $t\bar{t}$ purity is found when E_T^{miss} >85 GeV . Vertical scale is proportional to background absolute rate.

The invariant mass distribution of the lepton pair obtained in this way is the one in Fig. 4.5 (left) where, in the lower plot, the vertical scale of the two stacked distributions indicates the number of events populating the control region for 10 fb^{-1}. In order to maximize the statistic power, the control region is lower limited to 50 GeV/c^2 (to avoid possible contributions from low energy process, *e.g.* from some QCD combination) and pushed up to higher possible values where some $t\bar{t}$ events may enter.

The signal (upper plot) is here shown just for comparison but it gives no contributions in the control region. Since background is enhanced (and signal reduced by the missing energy cut), it can be calculated that signal would contaminate the background for more than 0.1% only if the FCNC branching ratio would be greater than ~ 0.15, that is clearly not the case.

The further step is to assess an optimized region around the peak in this background distribution, such to minimize the statistical error and the $Zb\bar{b}$ contamination. It has been accomplished by adapting a polynomial fit to the left side band of the $M(ll)$ peak (where statistic is much larger), then evaluating the number of $t\bar{t}$ events by integrating the result of the fit. A 3rd degree polynomial has shown to have a good convergence. This number of events is affected by two important uncertainties:

- statistical uncertainty of Poissonian type, that has mentioned before;
- contamination from the other background sources, here only the $Zb\bar{b}$. Possible contaminations from other sources are discussed below.

It is worth to observe that these $Zb\bar{b}$ residual events are not subtracted from the $N(t\bar{t})^{Contr}$, since their number is estimated from Monte Carlo technique, thus it would re-introduce the

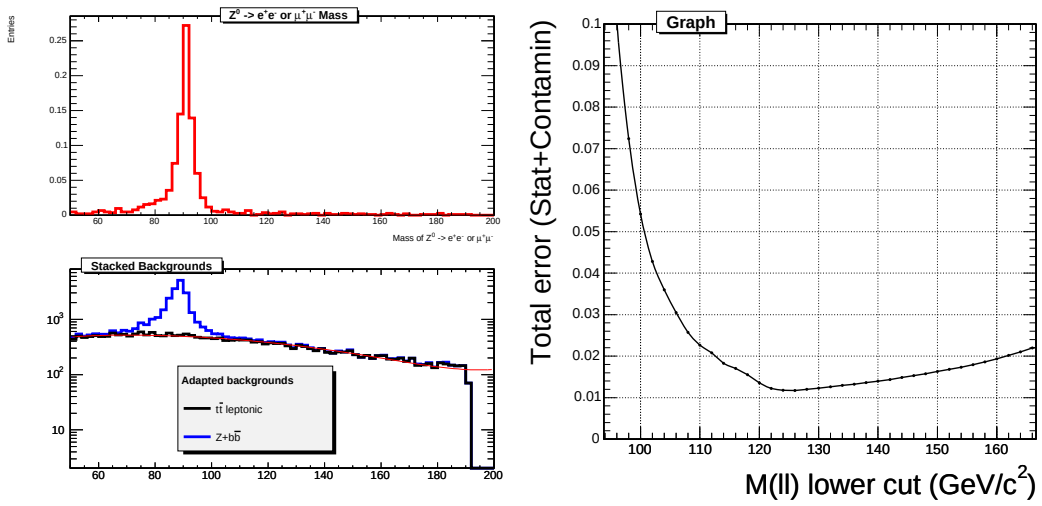

Figure 4.5: Left: Invariant mass of two same-sign opposite-flavour leptons, performed in the control region defined in the text. The $Zb\bar{b}$ background, that is showed stacked on $t\bar{t}$, gives the minimal contribution outside the $M(ll)$ peak. Contribution of the signal in that region (upper plot) is totally negligible. A polynomial fit is superimposed to the $t\bar{t}$ background. Right: the total error (statistic+contamination) in the counting of $t\bar{t}$ events, as a function of the lower integration point in the $M(ll)$ off-peak distribution. Optimal point is reached when integrating from $125\,\mathrm{GeV}/c^2$.

related uncertainty problems. It is considered just as an additional (little) systematic effect and added in quadrature.

The total error as a function of the lower integration limit in the $M(l^+l^-)$ is presented in Fig. 4.5 (right). The increase at low energy is clearly imputable to the approaching to the Z peak of the $Zb\bar{b}$ background, while the slow rise at high energy is due to the lack of event when narrowing the control region. The error is minimized at the optimal value for the lower cut $M(l^+l^-) = 125\,\mathrm{GeV}/c^2$. Here the total relative error on the number of event is 0.011, completely dominated by the statistical error.

Using the parameters determined from this fit, the number of $t\bar{t}$ events inside the mass peak can be estimated. The same mass window optimized for the signal selection is adopted here. When the rates are scaled to the $10\,\mathrm{fb}^{-1}$ integrated luminosity, the value $N(t\bar{t})^{Contr} = 8374 \pm 290$ $(10\,\mathrm{fb}^{-1})$ is found, where the error is given by the fit.

If the integrated luminosity is pushed to $100\,\mathrm{fb}^{-1}$, a different slope in the statistical error is originated and the fit can be extended down to $M(l^+l^-) = 102\,\mathrm{GeV}/c^2$, where a 0.0029 precision is reached. Using the fit determined in this region, the value $N(t\bar{t})^{Contr} = 55060 \pm 833$ $(100\,\mathrm{fb}^{-1})$ is found.

Following the Eq. 4.1, the events number in the control region has to be multiplied for the $\epsilon_j^{C\to S}$ factor, in order to be adapted to the signal analysis. For this purpose, all the requirements from the control region have to be scaled back in the signal region that has been optimized in the previous chapter. Therefore, the efficiency factor must include two

contributions:

- the ratio between the amount of missing energy in the signal region (E_T^{miss} >35 GeV) and in the control region (E_T^{miss} >85 GeV). At 10 fb^{-1}, it is estimate to be 1.342 $\pm$ 0.011(stat.) $\pm$ 0.040(syst.), where the systematic uncertainty comes from the effect on E_T^{miss} resolution discussed in the previous section. When higher luminosities will be integrated, statistical error will be naturally reduced and systematic effect will be limited as detailed above. This factor results $1.342 \pm 0.003(\text{stat.}) \pm 0.013(\text{syst.})$ after 100 fb^{-1};

- the effect of subsequent selections designed for the signal. This corresponds to the cumulative efficiency of all the cuts from the Z selection to the FCNC top invariant mass determination, but corrected for the presence of missing transverse energy and one b-jet. At 10 fb^{-1}, it is estimate to be $(2.01 \pm 0.14(\text{stat.}) \pm 0.30(\text{det. syst.}) \pm 0.23(\text{th. syst.})) \times 10^{-3}$, where the detector systematic errors are related to lepton efficiency and light-jet request. Uncertainty from theoretical errors (mostly from quark fragmentation) are quoted separately, because they will be treated differently in the significance estimation. The large statistical error is due to the lack of $t\bar{t}$ events in the last selection, even though selection efficiency has been evaluated in a subsample enriched in background content – that has been produced by enlarging the Z mass window. For higher integrated luminosity, systematic error will turn out to be dominant, since the factor results $(2.01 \pm 0.04(\text{stat.}) \pm 0.10(\text{det. syst.}) \pm 0.18(\text{th. syst.})) \times 10^{-3}$ after 100 fb^{-1}.

It is important to emphasize that the control region is built exploiting almost only the leptons kinematic features, thus it has no dependence from jet selection strategies. Also, the quantity $M(ll)$ is a Lorentz invariant so unaffected from recoil of jets or neutrinos. This implies that the $t\bar{t}$ shape, on which the first of the listed items is based, is fully independent from the other factor: as a result, the two contributions can be factorized and their errors added quadratically. The result of this operation provides these final estimates:

$$\begin{aligned} N(t\bar{t})^{Sign} &= 18.1 \pm 1.4(\text{stat.}) \pm 2.7(\text{det. syst.}) \pm 2.0(\text{th. syst.}) \quad 10\,\text{fb}^{-1} \\ N(t\bar{t})^{Sign} &= 149 \pm 4(\text{stat.}) \pm 7(\text{det. syst.}) \pm 13(\text{th. syst.}) \quad 100\,\text{fb}^{-1}. \end{aligned} \tag{4.2}$$

A further bonus of the approach is that, in principle, this estimate includes other additional higher order processes that come together with the tree-level $t\bar{t}$ production, as $t\bar{t} + b\bar{b}$ or $t\bar{t} + nj$. As they are expected to share the same shape with the $M(ll)$ distribution, there is no need to separately account for them.

As a final comment, it is worth to say that any other contaminating effect in the designed control region should be only imputable to some X resonance with 100< $M(X)$ <250 GeV/c^2 , produced with a cross section comparable to $t\bar{t}$ and leading to a $b + 2l + E_T^{miss}$ final state. Among the possible new physics scenario, the branching ratio for a $t\bar{t}H$ production is too

little to be significant, but some marginal role might be played by $b\bar{b}H/A$ production (with a SM or MSSM Higgs decaying in a final state with leptons and large missing energy, as $H \to WW$ or $A \to Zh$), as well as some tricky path of SUSY particles decay (*e.g* a stop decay $\tilde{t} \to t\tilde{\chi}_2^0 \to tl\tilde{l}_R \to tll\chi_1^0$). Adopting a conservative attitude to the sensitivity of the selection to new physics hints, none of these signals is considered in the simulation. On the other hand, counting in this window will tell the final world: if some 'bump' will be found somewhere, it will be consistently subtracted from the continuous background.

Estimation of the $Z + b\bar{b}$ background

Most of the advantages of choosing the Z mass distribution, explained above, can be exported even to the estimation of $Zb\bar{b}$ from data. The idea is to use the complementary region, *i.e.* the Z peak in the same $\Delta M(ll)$ defined for the signal, and simply count inside it to evaluate the amount of background. The follow additional considerations hold:

- the request for more than one b-tagged jet is mandatory, otherwise different $W/Z + j$ (where j indicates one or more jets, light or b-tagged) would populate the mass window, due to their large cross sections;
- in the previous chapter, b-tagging efficiency has proven to be some 20% lower than for $t\bar{t}$ processes and tagging two or more jets is more unlike. Even if the rate of $Zb\bar{b}$ is quite high, there is a strong reduction after the double b-jet requirement, thus a further lepton cannot be asked without a dramatic decrease in statistic;
- as most of $Zb\bar{b}$ events have a missing transverse energy lower than tens of GeV, the threshold on E_T^{miss} has been lowered to 20 GeV. This turns out to be helpful in recover a part of $Zb\bar{b}$ events; on the other hand, the strong request for a double-tagging prevents any other boson+jets process to contaminate the region;
- the $t\bar{t}$ background is demonstrated to have a contribution in the $\Delta M(ll)$ window. This is not a problem anymore, since in the previous section it has been demonstrated that the number of these events can be effectively estimated, by extrapolating a side band to the window;
- similarly to the previous analysis, the top FCNC signal, here not yet "discovered" by other cuts, has a totally negligible impact on the selection;
- some process not yet seen, as a $b\bar{b} + H/A$ production followed by a $H \to ZZ/WW$ or $A \to Zh$, might fall in that control region. Since here any shift along the invariant mass region is performed (as the $\Delta M(ll)$ for the signal is the same), including these processes would be an automatic effect of counting in this region.

In summary, the presence of a reconstructed Z in an invariant mass plot, supplemented by the request for at least two b-jets, is considered efficient in discarding all other boson+jets

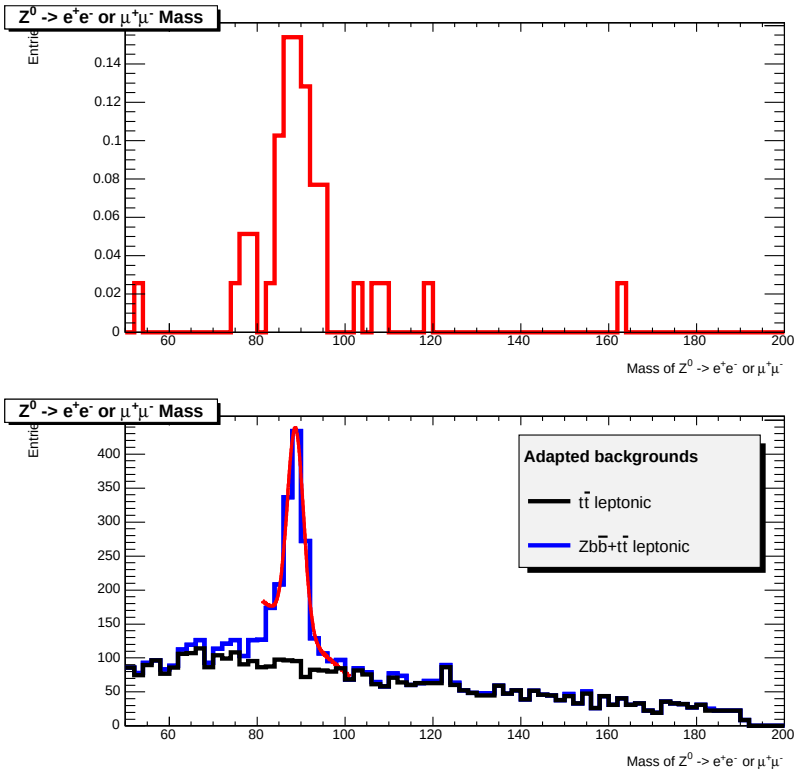

Figure 4.6: The distribution of lepton pair invariant mass in the region for the control of the $Zb\bar{b}$ background defined in the text. Contributions from other sources in the region are negligible and the $t\bar{t}$ background can be subtracted. A polynomial plus Gaussian fit is overlaid in the region where signal has to be searched.

contribution (included the tricky $W+b\bar{b}$, $Z+bj$ or single top). Therefore, the estimate of the $Zb\bar{b}$ from data starts from counting in the $\Delta M(ll) = 10\,\mathrm{GeV}/c^2$ window, where the amount of $t\bar{t} \to 2l$ events is now available from data. The invariant mass distribution is presented in Fig. 4.6 where, in the both plots, the vertical scale indicates the number of events populating the control region for $10\,\mathrm{fb}^{-1}$.

No optimization of the integration region is required here, since the counting will be directly performed in the $10\,\mathrm{GeV}/c^2$ window around the nominal Z mass. A Gaussian plus a polynomial is found appropriate to fit the total background under the peak: consequently, the count sums up the contribution under the peak plus a bias. The total number of count – that is supposed to include the $Zb\bar{b}$ component stacked on top of $t\bar{t}$ – amounts to $N(Zb\bar{b}+t\bar{t})^{Contr} = 1864 \pm 44$ and the error is the Poissonian statistic plus the error on fit. The former components dominates, and is expected to scale with the square root of the number of events. To clean the sample from the $t\bar{t} \to 2l$ events, the polynomial fit is used to evaluate the bias under the peak in Fig. 4.6 that, when integrated inside the 10 GeV windows, reads $N(t\bar{t})^{Contr} = 960 \pm 38$. From this method, the number of $Zb\bar{b}$ in the control region is obtained from subtraction as $N(Zb\bar{b})^{Contr} = 964 \pm 58$.

When this estimated number is scaled back to the signal region via the $\epsilon_j^{C\to S}$ factor, the effects to be included are:

- the ratio between events with only one b-jet (as requested for the signal region) and with more than one b-jet (as requested from the control region). It suffers from the systematic uncertainties on the fraction of the single-tagged events, and it turns out to be $9.3 \pm 0.5(\text{stat.}) \pm 0.8(\text{syst.})$ after $10\,\text{fb}^{-1}$ and $9.35 \pm 0.15(\text{stat.}) \pm 0.65(\text{syst.})$ after $100\,\text{fb}^{-1}$;

- the reduction in the amount of missing energy accepted for the signal region, *i.e.* $\text{N}(E_T^{miss} > 35\,\text{GeV})/\text{N}(E_T^{miss} > 20\,\text{GeV})$, that is estimated as $0.5117 \pm 0.0014(\text{stat.})$ at low luminosity. Each k-fold increase of integrated luminosity will result in an improvement of the error by a factor $\sqrt{k}$. Systematic uncertainties ranges from 3% at $10\,\text{fb}^{-1}$ to 1% for more than $\sim 50\,\text{fb}^{-1}$;

- the effect of subsequent selections designed for the signal, as the third lepton, missing energy cut, $M_T(W)$ and $M_T(bW)$ constraints, light-jet optimization. As in the previous case, it corresponds to the set of cuts from the Z identification down to the final selection, once subtracted for the efficiency of at least one b-tagged jet and the presence of missing energy. It results $(1.8 \pm 0.8(\text{stat.}) \pm 0.3(\text{det. syst.}) \pm 0.2(\text{th. syst.})) \times 10^{-4}$ after $10\,\text{fb}^{-1}$, and $(1.8 \pm 0.3(\text{stat.}) \pm 0.1(\text{det. syst.}) \pm 0.1(\text{th. syst.})) \times 10^{-4}$ after $100\,\text{fb}^{-1}$.

In principle, one cannot assume these contributions independent: in fact, b quark decays play a role in producing missing energy, so a different choice of the E_T^{miss} may lead to a different distribution in the b-jet multiplicity. This point has already be discussed when the b-jet selection was exploited with the missing energy requirement: though some negative correlation is not excluded, it is below the precision of some percent of this analysis. Therefore, systematic uncertainty on the missing transverse energy is quadratically summed to that on single-tagged fraction.

The result of this operation provides these final estimates:

$$\begin{aligned} N(Zb\bar{b})^{Sign} &= 1.7 \pm 0.7(\text{stat.}) \pm 0.3(\text{det. syst.}) \pm 0.2(\text{th. syst.}) \quad 10\,\text{fb}^{-1} \\ N(Zb\bar{b})^{Sign} &= 16.6 \pm 2.3(\text{stat.}) \pm 1.4(\text{det. syst.}) \pm 1.5(\text{th. syst.}) \quad 100\,\text{fb}^{-1}. \end{aligned} \tag{4.3}$$

These estimates are significantly dominated by the large statistic error, induced from the very little number of events in the signal region – even if cuts softening is applied to evaluate partial efficiencies. Nevertheless, this is not a major issue on the total background error, as $Zb\bar{b}$ is a factor ten weaker than $t\bar{t}$.

Both background sources are now supposed to be estimated to such well-suited control regions. Calculating the total background plus its errors is straightforward:

$$N(TOT)^{Sign} = 19.8 \pm 1.6(\text{stat.}) \pm 2.7(\text{det. syst.}) \pm 2.0(\text{th. syst.}) \quad 10\,\text{fb}^{-1} \qquad (4.4)$$
$$N(TOT)^{Sign} = 166 \pm 5(\text{stat.}) \pm 9(\text{det. syst.}) \pm 13(\text{th. syst.}) \quad 100\,\text{fb}^{-1}.$$

In the next section, $N(TOT)^{Sign}$ with its uncertainties will be combined with the $t \to qZ$ signal efficiency, in order to produce reasonable upper limits for the FCNC branching ratio, that is the ultimate goal.

4.2.3 Background estimation for the $t \to \gamma q$ channel

The determination of the rate of background that is relevant for the $t \to \gamma q$ analysis presents some more challenges, and is difficult to perform in the same clean way. Referring to results presented in the previous chapter, some observations can be drained:

- as outlined in the signal selection phase, no signature as distinctive as the multi-lepton exists here, so a cumulative background from diverse sources have chances to survive down to the very last cuts. Namely, the W/Z+jet background, having large cross sections even in the addressed kinematic region, can be suppressed only with rather hard requirements;
- the processes that at the end of selection turn out to be most relevant, *i.e.* the inclusive $t\bar{t}$ production and the single-top in the t-channel, share a very similar behaviour because of the presence of standard decaying top quark in both samples. Separation of the two sources is not straightforward and has to rely on kinematic cuts and number of reconstructed jets;
- in the very last cut, single-top contribution has resulted suppressed down to a very low level: in the signal region, single-top rate is expected to be lower than $t\bar{t}$ by almost two orders of magnitude.

The last feature motivates the assumption to consider the single-top not a relevant background, so to not estimate it from data. Only about 0.6 events are expected for each $10\,\text{fb}^{-1}$ of data and, even considering the large systematic error related to the estimate, it would be completely drowned by the total error on the top pair background. Therefore, only the choice for the control regions in $t\bar{t}$ is addressed in the following.

Estimation of the $t\bar{t} \to l + X$ background

The results discussed in the analysis for the top photonic channel clearly show that the region where the FCNC signal is determined, *i.e.* the $M(q\gamma)$ invariant mass, is not suitable to estimate the background since the number of surviving $t\bar{t}$, though larger than single top,

is still poor. Hence the $M_T(bW)$ invariant mass is addressed, once properly enriched with top pair contents.

Some among the observations raised in that analysis suggest the criteria to enrich the control region. The most useful ones have found to be the following:

- transverse momentum of the reconstructed W above a threshold. Referring to Fig. 3.62 in the previous chapter and discussion therein, this variable demonstrates helpful in partially decouple the background. The best value of this threshold will be set below, with the aim to minimize the total error;
- a veto for less than two light-jets. In the discussion on the number of jets and from distribution in Fig. 3.66, it has been observed that the number of events with more than one reconstructed jet is more than an order of magnitude larger in $t\bar{t}$ than in single-top. In fact, since the former includes all possible final states resulting in a hard lepton, thus with a large hadronic activity, the latter has been considered relevant in the only W leptonic decay, so at tree level the only jet is expected to be the one accompanying the top – that often fails detection. Requiring the existence of some light-jets already at this level has the bonus to not introduce the systematic uncertainty in the jet energy scale. Since jets are already included in the control region, that uncertainty – that has demonstrated to hardly affect the $t \to qZ$ analysis – is no more an issue here;
- since the multiplicity of b-tagged jets may suffer from different systematic effects, no discrimination based on the b-jets number is adopted, and the request to veto a second jet – that was made for the signal – is left untouched.

This region is displayed in Fig. 4.7 where, in addition to the target background sources, also the W+jets contribution is shown. The request for a large boost and a hard jets multiplicity rules out such contribution. Therefore, at this level the only handle is the optimization of the upper cut in the $p_T(l\text{-}E_T^{miss})$ variable. The minimum of the total error, given by the statistical uncertainty summed in quadrature to the contamination from single-top and W+jets, will dictate this threshold value.

The total number of event has found by the integration of a Gaussian plus a fourth degree polynomial. The best point is found when taking $p_T(l\text{-}E_T^{miss}) > 85\,\mathrm{GeV}/c$. At such optimal threshold, contamination from single-top is 3.2%; when the statistical error is added, a total 3.9% uncertaity is obtained. Then the estimate from data of top pair inclusive production turns out to be $N(t\bar{t})^{Contr} = 1738 \pm 68$.

The rescaling of this estimate to the signal region clones the signal selection for the FCNC top and removes the background enhancement factors. That is, the $\epsilon_j^{C\to S}$ number has to account for:

- the different W transverse momentum thresholds, that are not optimized to the same upper cut. The ratio between the fraction of events taken with the cut for the signal

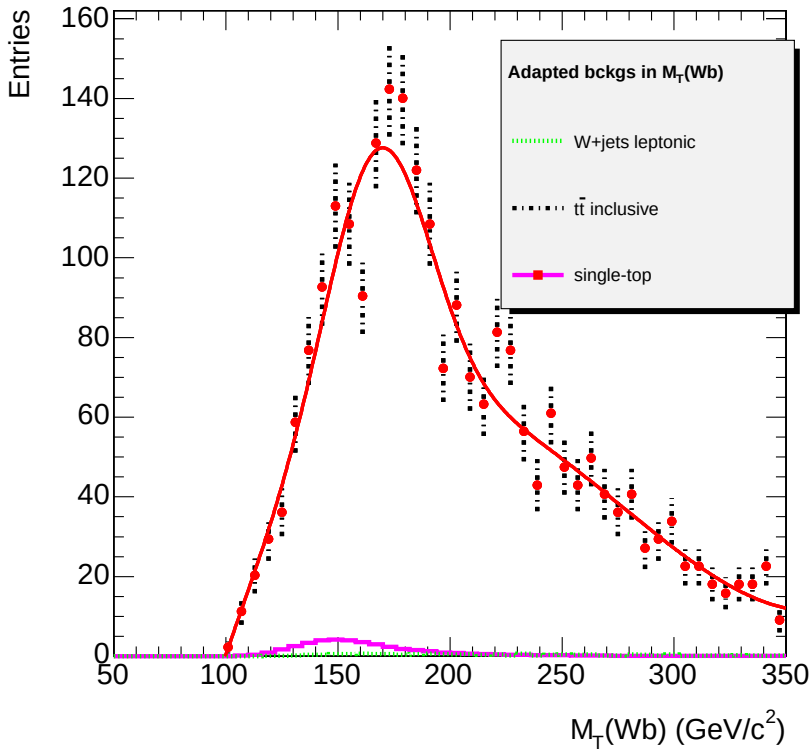

Figure 4.7: The distribution of the transverse mass for the top decaying in W and b-jets, once enriched in top pair content as explained in the text. The W+jets process (where the jet rises from parton with $85 < \hat{p}_T < 150\,\mathrm{GeV}/c$) though has large cross section is almost completely erased from this window, while single-top is no more than 3%. Contributions are showed superimposed (and not stacked) and a Gaussian plus a fourth degree polinomial fits the distribution.

region, over the fraction of events taken with the cut for the control region is estimated to be $1.488 \pm \delta x$, where the statistical error is $\delta x = 0.046$ for $10\,\text{fb}^{-1}$ and $\delta x = 0.015$ for $100\,\text{fb}^{-1}$. The systematic error derives only from uncertainties on the lepton resolution and momentum scale, along with E_T^{miss} resolution, and it is 0.065;

- the fact that no vetoes on the number of light-jet is imposed on a signal. Therefore, rescaling has to be done for the inverse of the frequency for a more-than-one light-jet. This number receives systematic effects both from detector (induced by b-tagging and JES) and from theoretical uncertainties in light quark fragmentation. It reads $1.51 \pm 0.10(\text{stat.}) \pm 0.01(\text{det. syst.}) \pm 0.15(\text{th. syst.})$ after $10\,\text{fb}^{-1}$ and $1.51 \pm 0.03(\text{stat.}) \pm 0.005(\text{det. syst.}) \pm 0.15(\text{th. syst.})$ after $100\,\text{fb}^{-1}$;

- the reconstruction of the FCNC top with the $t \to \gamma q$, that adds to the selection the request for the hard photon and the invariant mass bound. This efficiency is estimated by the last cut in the signal region, once correction is made for the efficiency of the hard jet – that is already supposed to be present in the sample. The only systematic effect to be included for this last term is related to the photon, that basically suffers only from the ECAL energy scale uncertainty (EES). This value is $(8.51 \pm 1.13) \cdot 10^{-3}$, at $10\,\text{fb}^{-1}$; statistical error smoothly decreases with integrated luminosity, while EES from Tab. 4.3 is only a little 0.5%.

Putting together the factors corresponds to find what is assumed to be the total background in the signal region. Such values are found at the different luminosities:

$$\begin{aligned} N(TOT)^{Sign} &= 33.2 \pm 4.6(\text{stat.}) \pm 1.1(\text{det. syst.}) \pm 3.7(\text{th. syst.}) \quad 10\,\text{fb}^{-1} \\ N(TOT)^{Sign} &= 332 \pm 15(\text{stat.}) \pm 11(\text{det. syst.}) \pm 37(\text{th. syst.}) \quad 100\,\text{fb}^{-1}. \end{aligned} \tag{4.5}$$

A degree of speculation is present also in this control region, since processes different from the Standard Model – thus not surely ruled out by the analysis – can enter the window. A possible new physics signature between 50 and $350\,\text{GeV}/c^2$ should result in a b-jet and in a W with some boost, accompanied by a significant hadronic activity and capable to pass a photon trigger. Decay channels for the Higgs boson (both SM and MSSM) could provide such signatures, manifesting in resonance in the spectrum (as for instance $H^+ \to t\bar{b} \to Wb\bar{b}$) or in an excess on the continuum (as the $b\bar{b}H$ production followed by a $H \to WW$). Though the photon trigger should do a good job in rejecting some of these processes, the possibility to observe something new in this control region is still an issue. Instead of claim for 'contaminations' in the selections, this could be regarded as an interesting side-effect of the analysis.

4.3 Estimation of statistical significance

Following the common approach, the estimation of statistical significance consists in comparing the 'null hypothesis' H_0, stating that only Standard Model is at work and no top anomalous decay is present, and an alternative hypothesis H_1, which states the presence of a FCNC signal on top of the background. The alternative hypothesis is related to the branching ratio of the considered decays, that in turn depends from top anomalous couplings, which are to be determined as the ultimate goal. The invariant mass distribution of the qZ/γ final state depends on the true state of nature being H_0 or H_1.

4.3.1 Significance definitions

The term significance usually refers the "number of standard deviations" an observed signal is above expected background fluctuations, assuming the signal follows a Gaussian distribution with a standard deviation of one.

As a consequence, a given value of significance corresponds to the probability that the claimed signal is caused merely by fluctuations of the background, and this probability is obtained by performing the corresponding integrals of a standard Gaussian distribution. Since a signal is usually searched for in many bins of a distribution, a very high value of the significance must be used before an observed "peak" found somewhere can be declared to be a signal observation. For asymmetric distributions, the significance can be quoted in terms of equivalent standard deviations of the normal Gaussian.

In the following, sensitivity estimates will be done for the "discovery claim" at 5σ and the "strong evidence" at 3σ. When the significance is 5σ, the corresponding one-sided Gaussian probability that a local fluctuation of the background mimics a signal is $2.9 \cdot 10^{-7}$.

Several methods exist to quantify the statistical significance of an expected signal at future experiments. For the present purpose, the focus has to be put on the problem of a 'counting experiment' in presence of background, where the significance depends of the number of signal events N_S observed in some signal region of a statistical distribution, and the number of expected background N_B, which is assumed to be Poisson-distributed. This procedure requires working with binned distributions, which in turn means that bin positions and bin widths have to be fixed. The common significance definitions are:

- $S_{c1} = N_S/\sqrt{N_B}$. This expression is a true estimation of the significance only in the high statistic limit, *i.e.* when the N_B distribution is strictly Gaussian. It has been shown that S_{c1} is suitable only for background levels larger than some 50 events, otherwise it badly overestimates the sensitivity at low backgrounds;

- $S_{c2} = N_S/\sqrt{N_B + N_S}$ is true even for a small numbers of events following a Poisson distribution. For high significance settings, it is expected to slightly overestimate the sensitivity at low backgrounds;

- $S_{c12} = 2\left(\sqrt{N_B + N_S} - \sqrt{N_B}\right)$ has been suggested in literature [143], that produces an intermediate solution. The formula for S_{c12} is strictly only valid in the Gaussian limit, but tabulated values exist for small statistics. This significance definition has been established as a common reference for all the CMS Physics Group.

4.3.2 Significance estimation

In the event counting approach adopted here, the input parameters to evaluate the experiment sensitivity are the number of signal and background events provided by the analysis. It is implicitly assumed that the N_S, N_B obtained coincide with the average number of their distributions. In a simple counting experiment, the observable N is the number of observed counts. Null hypothesis H_0 is defined as the distribution of N being a Poisson with the mean equal to N_B, while alternative hypothesis H_1 is that the distribution is instead a Poisson with a larger mean $N_B + N_S$. Hence a minimum number N_0 of events has to be found, as the discriminating value between the two hypothesis. That value can be found via an input parameter indicating the probability of the discovery, or alternatively via the confidence limit in the Standard Model predictions if no signal will be found.

As stated above, a pre-defined significance level a is related to the the probability to find number of selected events over a certain pre-specified critical value, beyond which the validity of H_1 is assumed. This implies that the probability of the background fluctuations over a $N > N_0$ is equal to a pre-specified value α, *i.e.* N_0 is such as:

$$P(N > N_0|H_0) = \alpha, \quad \text{where } P(N > N_0|H_0) = \sum_{N_0}^{\infty} f(N|N_B), \tag{4.6}$$

and $f(N|N_B)$ is the Poisson distribution for N with mean N_B, $f(N|N_B) = e^{-N_B} N_B^N / N!$. If $N_0 = N_S + N_B$ is assigned, α is the probability to have a background fluctuation emulating a signal, resulting in accepting the H_1 when H_0 is actually true (called 'Type II error'). The probability α is usually accommodate to a 'tiny' value and the significance level a is obtained converting this probability into an equivalent number of a normalized Gaussian sigmas. Therefore:

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_a^{\infty} e^{-x^2/2}. \tag{4.7}$$

In other words, if a potential observation found no evidence for a deviation from the standard theory, the 'confidence level' of this negative result (C.L.) is given by α or a.

The other element to be considered in a test is the probability that a discovery is made. This calls for the probability β of counting $N < N_0$ in a model with new physics, that is:

$$P(N < N_0|H_1) = \beta, \quad \text{where } P(N < N_0|H_1) = \sum_{0}^{N_0} f(N|N_B + N_S), \tag{4.8}$$

and $f(N|N_B)$ is the Poisson distribution for N with mean $N_B + N_S$, $f(N|N_B + N_S) = e^{-(N_B+N_S)}(N_B + N_S)^N/N!$. If $N_0 = N_S + N_B$ is assigned, β is the probability to have a signal fluctuation emulating a background, resulting in rejecting the H_1 when it is actually true (called 'Type I error'). The complementary part $1-\beta$ is the probability that a discovery will be claimed assuming H_1 is true ("power function") and it is clearly desirable to maximize it.

Even the β value can be converted into an equivalent number of a normalized Gaussian sigmas, as

$$\beta = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b} e^{-x^2/2}. \tag{4.9}$$

In the present study, both α and β depend from the magnitude of the anomalous decay coupling leading to the searched signal.

In the special case $N_B \gg N_S$, the background can be assumed to obey the normal distribution. It can be shown that when $\beta = 0.5$ is assumed, the discovery significance takes the simple form S_{c2} indicated above, that for even larger background approaches to the definition of S_{c1}. When the hypothesis $N_B \gg N_S$ is no longer valid, it is shown that an attractive estimator of the discovery significance could be the S_{12} [144], with the assumption that signal and background still follow the normal distribution. In the case of the asymmetrically distributed signal and background, these significances can be used only as approximations.

If a stronger statistical power is desired, the critical N_0 has to be lowered, hence the α probability from Eq. 4.6 will be larger, and the confidence level C.L. can decrease. In that case, the significance estimator is re-formulated as

$$S_{12}^{*} = 2\left(\sqrt{N_B + N_S} - \sqrt{N_B}\right) - k(\beta). \tag{4.10}$$

If $1-\beta$ is intended as the one-side probability from Eq. 4.8, it is easy to interpret $k(\beta)$ as the distance of the critical N_0 (in number of sigmas) from the mean of the signal distribution, assuming it to be Gaussian. The $k(\beta)$ function is tabulated and reads, for instance, $k(\alpha = 90\%) = 1.28$ and $k(\beta = 95\%) = 1.64$.

In the following, the analysis of results will be presented for a given significance and Confidence Level C.L. . Following an idea outlined in Ref. [146], the confidence band for anomalous coupling will be defined by building a region where the power of the test is greater or equal to the Confidence Level, as $1-\beta >$C.L.. This specific choice is particularly informative, because it defines the region in the parameter space for which the experiment will certainly give an answer. If performing the experiment does not lead to discovery, the resulting limits will exclude that region at the chosen C.L., and if a discovery will be claimed with a chosen significance, it has a probability at least equal to C.L.. Moreover, it is independent of *a-priori* expectations about the presence of a signal in new phenomena and, as explained later, clearly indicates the parameter space that can be optimized.

4.3.3 Inclusion of systematic uncertainties

All systematic effects are supposed to have an impact on new physics discoveries, because they affect the amount of background events and their distribution, thus the control region is modified. The estimation of relevant background rate from data is an essential tool to reduce effectively the systematic fluctuations, and results in a new physics sensitivity that is robust against experimental and theoretical variations. While general strategies have been designed to incorporate systematic uncertainties in significance and confidence level calculation [143], a conceptually simple way is to consider the probability for the observed number N as a conditional Poisson probability, where the mean N_B may systematically fluctuate. This can be done by numerically convoluting the Poisson distribution (or the Gaussian distribution in the high-statistics limit) with the probability density function of the theoretical uncertainty. As a result, possible variants of significance definitions including systematic errors can be provided.

The most simple assumptions are considered for the probability density functions of systematic effects:

- systematic theoretical uncertainties, which in principle can feature any dependence from input parameters that is hard to know precisely, are assumed to follow a uniform distribution with a ΔN_B^{THEO} standard deviation;

- systematic experimental uncertainties, that come from several measure errors each having a 'small' effect, are assumed to follow a Gaussian distribution with a ΔN_B^{EXP} standard deviation.

Since the procedure outlined above has showed that, once the background is estimated in some suitable signal region, all systematic effects are limited to few percent, this assumptions are completely motivated.

The numerical implementation of such systematic variation having these statistical properties is elaborated by a simple program developed by S. I. Bityukov [145] that has been widely used by the CMS collaboration in the preparation of the Physics Technical Design Report. The program allows to calculate the significance S_{12} after having specified N_S, N_B and the total systematic effects ΔN_B. Theoretical errors are incorporated by a convolution with the probability density function assumed above. An approximated estimate of the S_{12} with systematic errors inclusion is given by:

$$S_{12}^{SYST} = 2\frac{N_B}{N_B + \Delta N_B^2}\left(\sqrt{N_B + N_S} - \sqrt{N_B}\right), \tag{4.11}$$

and the user can specify if he prefers this analytic approximation or a numerical implementation performed by a Monte Carlo simulation.

4.3.4 Sensitivity results of the analysis

Once a significance level for the result has been chosen, the statistical test $N > N_0$ allows to determine how much the obtained statistic is compatible with the H_1 hypothesis. In other words, a minimum number N_{min} of signal events has to be found that can produce the maximum power. This minimum number is related to the minimum branching ratio for FCNC decay that the analysis can detect, thus it can deliver the sensitivity reach for the anomalous decay.

Analysis cut optimization

The results presented above have been obtained with a set of cuts optimized for a maximization of the discovery significance and confidence level together. As mentioned above, the model proposed in Ref. [146] has been followed to accomplish the task.

Here the minimum signal N_{min} discriminating between the H_0 and H_1 hypothesis is related to the the Gaussian equivalent of the Poisson. The condition $1 - \beta >$C.L. can be expressed as

$$N_{min} = a\sqrt{N_B} + b\sqrt{N_B + N_{min}},$$

where a and b have been defined in Eqq. 4.7 and 4.9 as the number of sigmas corresponding to one-sided Gaussian tests.

The solution of the previous equation is:

$$N_{min} = \frac{b^2}{2} + a\sqrt{N_B} + \frac{b}{2}\sqrt{b^2 + 4a\sqrt{N_B} + 4N_B},$$

where both the minimum signal and the N_B depends from selection cuts. The best sensitivity will be obtained when the N_{min} is as small as possible.

As stated above, the FCNC branching ratio $BR(FCNC)$ is related to this N_{min} through the relation in Eq. 4.12. Expliciting the $BR(FCNC)$ from the equation, yields

$$BR(FCNC) = \frac{\frac{b^2}{2} + a\sqrt{N_B} + \frac{b}{2}\sqrt{b^2 + 4a\sqrt{N_B} + 4N_B}}{L \cdot \epsilon_S \sigma(t\bar{t}) \cdot BR}.$$

The aim of optimization is to minimizing this threshold visibility. Here the choice is to focus on the specific target case $a = b$ with a significance and confidence level corresponding to the 95%, though other reference values has been computed below. If this assumption is adopted, N_{min} shrinks to $a\left(a + 2\sqrt{N_B}\right)$ and the target becomes the maximization of the quantity $\epsilon / \left(a/2 + \sqrt{N_B}\right)$.

When high significance or confidence level are desired, it can be shown that tails of Poisson distribution significantly deviate from the Gaussian behaviour postulated here. Nevertheless, the Gaussian approximation can easily be improved, without loosing the good features of the

solutions, for instance performing an empirical fit of the Poisson shapes. This results in the following improved expression for N_{min}:

$$N_{min} = \frac{a^2}{8} + \frac{9b^2}{13} + a\sqrt{N_B} + \frac{b}{2}\sqrt{b^2 + 4a\sqrt{N_B} + 4N_B}.$$

In the case study with significance at 95% and power $1 - \beta = 0.5$, it turns out that the target is to minimize the quantity

$$\frac{\epsilon}{0.872 + \sqrt{N_B}}.$$

Therefore, all cuts in the analyses detailed in the previous chapter were aimed to maximize this ratio.

The general equation to be considered is:

$$N_{min} = L \cdot \epsilon_S \sigma(t\bar{t}) \cdot BR \cdot BR(FCNC), \tag{4.12}$$

where L is a given integrated luminosity, ϵ_S is the selection efficiency and $\sigma(t\bar{t})$ the production cross section for a pair of top quarks. The branching ratios $BR \cdot BR(FCNC)$ measures the probability for these quarks to lead to the final states that have been considered in the analyses. For the two decay channels investigated in the work, Eq. 4.12 splits in two parts:

$$N_{min}(t \to Zq) = 2 \cdot BR(t \to Zq) \cdot BR\,(W \to l\nu) \cdot BR\,(Z \to ll) \cdot \sigma(t\bar{t}) \cdot L \cdot \epsilon_S(t \to Zq)$$
$$N_{min}(t \to \gamma q) = 2 \cdot BR(t \to \gamma q) \cdot BR\,(W \to l\nu) \cdot \sigma(t\bar{t}) \cdot L \cdot \epsilon_S(t \to \gamma q),$$

where the leptonic branching fractions are intended for electrons and leptons.

In order to provide enough informations on the sensitivity of the analysis, as well to easily compare this results with past (and possibly future) estimates, three different confidence regions are defined:

1. significance at 95% and statistic power $1 - \beta = 0.5$. It has to be considered as a 'golden case' for which all selection cuts have been optimized, and can be directly compared with existing exclusion limits from HERA and Tevatron;

2. significance at 5σ level and confidence level at 90%. This can be assumed as the 'discovery' claim of a top flavour changing neutral current in those specific channels, and it is further enforced by a large confidence level;

3. significance at 3σ level and confidence level at 95%. It is an intermediate case, where the 'strong evidence' for the FCNC discovery is supplemented by a very high confidence level.

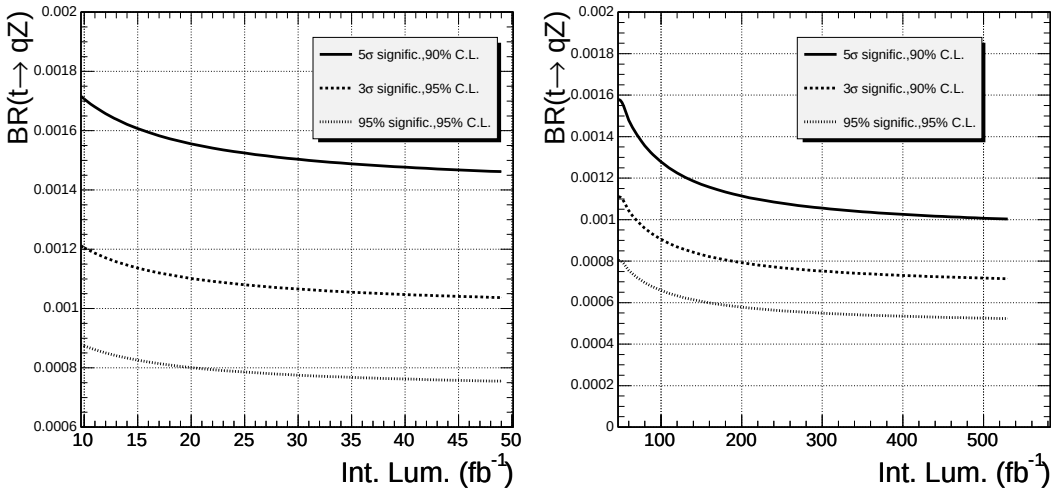

Figure 4.8: The upper limits for the $t \to qZ$ branching ratio, as expected in the 10-50 fb^{-1} (left) and in the the 50-500 fb^{-1} (left) range, for three different significance regions.

For each case, the value of N_{min} is estimated by the numerical implementation of the S_{12} significance estimation, with the proper corrections for the modified power earlier discussed. The systematics uncertainties are included with a standard distribution convolution.

Estimation of $BR(t \to qZ)$ and $BR(t \to q\gamma)$ sensitivity

At this level of analysis, all tools are in place to determine the number N_{min} of top FCNC decay, whence the maximum branching ratio to which the experiment is sensitive. Signal efficiency has been found in the previous chapter; background has been estimated from control region here above, along with its error; significance regions have been defined. The integrated luminosity is here considered as a running variable, and results are quoted in function of that. Since it has been demonstrated that selections can be nicely extrapolated up to the next LHC ages, upper limits till some hundreds of fb^{-1} are proposed. Branching ratios are evaluated with the numerical implementation of S_{12}^{SYST}, when adapted to the specified significance regions and with separately accounting for detector-related and theoretical systematic effects. The standard S_{c1}, with correction for systematics implemented, has been used only for the high luminosity situations where $N_B \gg N_S$.

The Fig. 4.8(left) represents the result for the initial low luminosity phase. It indicates that going from 10 to 50 inverse femtobarns is a major step, because it allows an improvement of the upper limits by some 10%, while some systematic errors could already be smaller than what has been considered here.

The Fig. 4.8(right) is the result of the extrapolation to the higher luminosity phase (10^{34}cm^{-2}s^{-1}), where all considerations about the impact of systematic uncertainties are took into account. Branching Ratio sensitivities is extended up to 500 fb^{-1}, assuming that no differences in the trigger and off-line selection parameters.

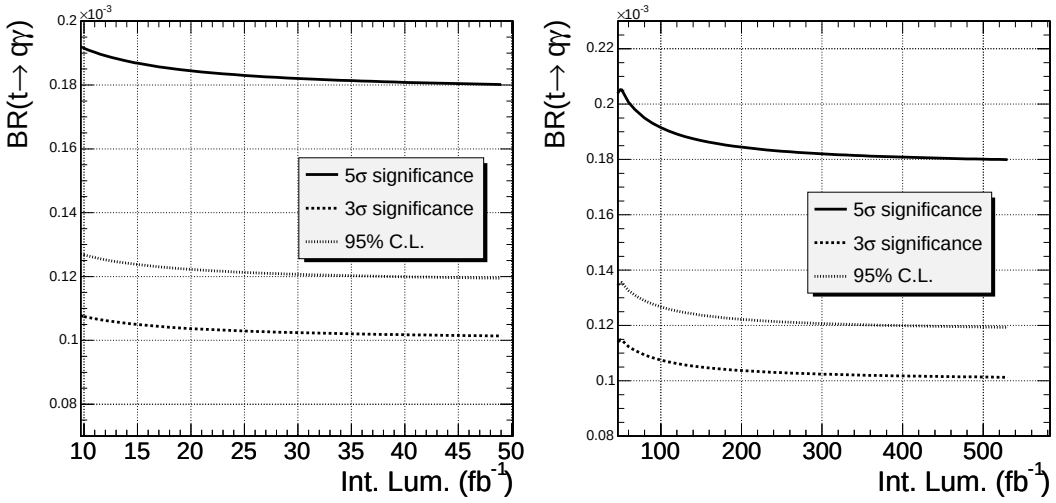

Figure 4.9: The upper limits for the $t \to q\gamma$ branching ratio, as expected in the 10-50 fb^{-1} (left) and 50-500 fb^{-1} range for three different significance regions.

Table 4.13: Results for the upper limits in the branching ratio for the $t \to qZ$ decay. The three different significance regions are indicated, for five reference integrated luminosities.

Int. Lumi (fb^{-1})	5σ, 90% C.L	3σ, 90% C.L	95% C.L
10	$17 \cdot 10^{-4}$	$12 \cdot 10^{-4}$	$8.7 \cdot 10^{-4}$
50	$15 \cdot 10^{-4}$	$10.4 \cdot 10^{-4}$	$7.6 \cdot 10^{-4}$
100	$13 \cdot 10^{-4}$	$9.1 \cdot 10^{-4}$	$6.6 \cdot 10^{-4}$
500	$10 \cdot 10^{-4}$	$7.2 \cdot 10^{-4}$	$5.2 \cdot 10^{-4}$

The same procedure has been set up for the photon channel, with results presented as a function of luminosity in Fig. 4.9(left) for the first phase and Fig. 4.9(right) for the next future. Obtain a better absolute limit in this channel is a common feature to all the analyses, basically because the photon has not to be scaled for the decay branching ratio as the Z does. On the other hand, the efficiency for the signal detection and the amount of surviving background between two channels have demonstrated to be fairly similar.

The dependence of the FCNC sensitivity to the amount of background is exemplified in Fig. 4.10, where the $t \to Zq$ case is analyzed. The integrated luminosity here is fixed as well as the signal efficiency, thus an enhancement of the background number of events B may come only from an underestimate cross section or instantaneous luminosity. Upper limits $BR(t \to qZ)$ at 95% C.L. are quoted as a function of B, for different contributions of systematic effects. The impact of the most important instrumental uncertainty, *i.e.* energy scale, is displayed on the left, while the effect of uncertainty on fragmentation model (that is the only theoretical issue that has some relevance) is represented on the right.

From these plots, it is clear that an improvement on sensitivity may come both from

Table 4.14: Results for the upper limits in the branching ratio for the $t \to q\gamma$ decay. The three different significance regions are indicated, for five reference integrated luminosities.

Int. Lumi (fb^{-1})	5σ, 90% C.L	3σ, 90% C.L	95% C.L
10	$2.51 \cdot 10^{-4}$	$1.76 \cdot 10^{-4}$	$1.27 \cdot 10^{-4}$
50	$2.35 \cdot 10^{-4}$	$1.65 \cdot 10^{-4}$	$1.19 \cdot 10^{-4}$
100	$2.13 \cdot 10^{-4}$	$1.50 \cdot 10^{-4}$	$1.09 \cdot 10^{-4}$
500	$1.95 \cdot 10^{-4}$	$1.38 \cdot 10^{-4}$	$0.992 \cdot 10^{-4}$

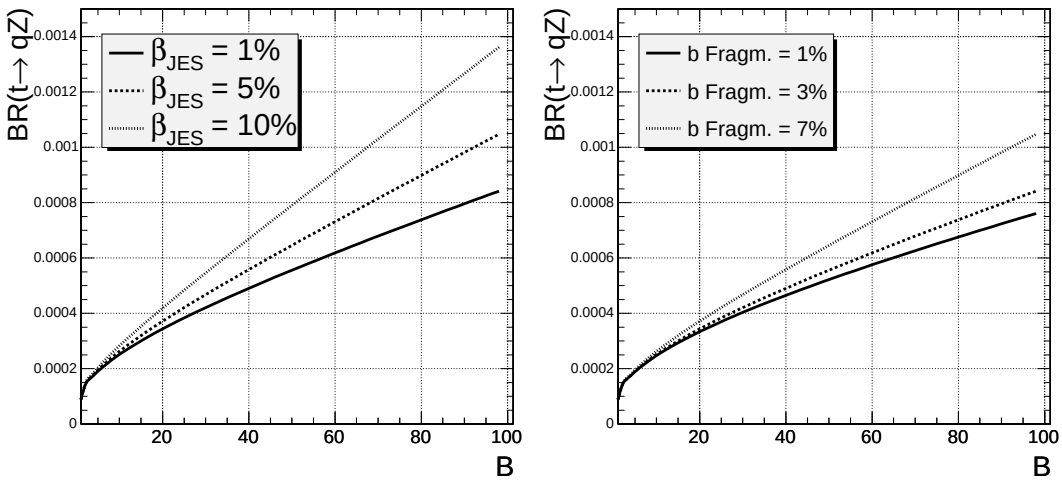

Figure 4.10: The sensitivity to FCNC Branching Ratio in the $t \to qZ$ channel, as a function of background event number with fixed instantaneous luminosity and signal efficiency. Three families of curves are represented for different values of jet energy scale uncertainty (left) and fragmentation uncertainty on b-jet parametrization (right).

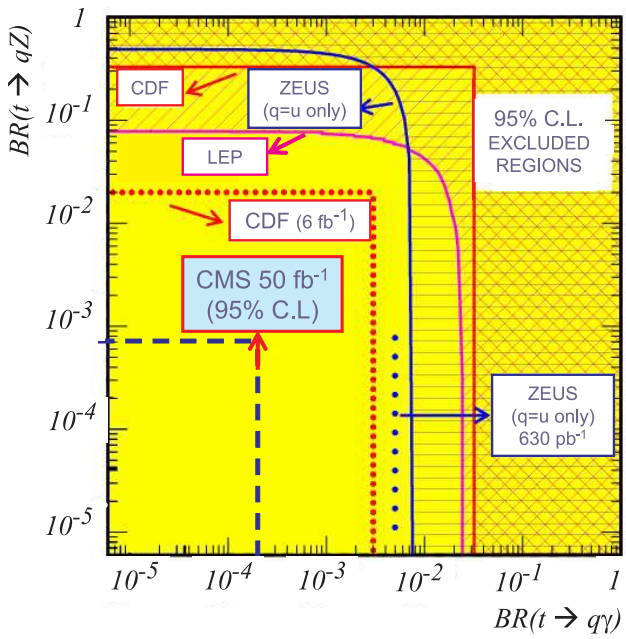

Figure 4.11: The exclusion plot with upper limits for the two FCNC addressed channels. The most important current limits have been inserted, from current and recent accelerators. CDF experiment is considered in the most optimistic scenario, using all the data at the end of Tevatron life and with a sensitivity improved by a factor two. Expectations for CMS exclusion limits, after $50\,\mathrm{fb}^{-1}$ are represented.

a better knowledge of the jet energy scale (that will be driven by data as more and more luminosity will be integrated) and from a tuning of hadronization models on data themselves. The significant dependance from B confirms that only an estimation of background from data is capable to precisely pin down the sensitivity of the analysis.

Comparison with current estimates

The advantages of using an apparatus as CMS for the search for FCNC, can be well appreciated by a comparison with current exclusion limits for anomalous couplings. The huge instantaneous luminosity and the large $t\bar{t}$ production rate, along with the brilliant performance of CMS detector and a careful optimization of the analysis, will allow remarkable improvements even in the very first years of LHC activity.

The Fig. 4.11 can now be compared with Fig. 1.7 pictured at the end of the first chapter. Here the moment when $50\,\mathrm{fb}^{-1}$ of integrated data will be available is taken for comparison. In this phase many systematic effects can supposed to be well controlled, and efforts will hopefully be applied to minimize the most annoying effects here, as b-tagging and jet energy scale uncertainties.

If – avoiding to be too optimistic – the possibility of a FCNC discovery is excluded to be observed by the analysis of current Tevatron and HERA data, CMS will soon start to put

upper limits with an unprecedent precision; similar behaviour is expected for ATLAS.

Anomalous coupling upper limits

The branching ratio upper limits quoted above could naturally lead to an exclusion plot for the anomalous couplings regulating the top FCNC decay. If one refers to Eqq. 1.7 in Chapter 1, deriving bounds on tZq and $t\gamma q$ vertexes are straightforward. Here a couple of assumptions are considered:

- the analysis sensitivities to $t \to uV$ and $t \to cV$ decays are the same. This is reasonable, since in performing the light-jet selection some efforts have done to equalize the response against the two flavours. Results presented in Sec. 3.2.3 (see for instance Fig. 3.11, left) give good confidence on that;
- no means are provided to separate the vectorial component in the $t \to qZ$ decay, $\Gamma(t \to qZ)_\gamma$ from the axial one $\Gamma(t \to qZ)_\sigma$. Theoretical studies do not seem to provide enough hints on this point, and limited resolution on most of kinematic variables would not allow to precisely discriminate between two modes. Therefore, here it is considered alternatively that the coupling with Z would be only axial or only vectorial, that is perfectly consistent with providing only the sensitivity 'upper limits'.

Using Eqq. 1.7 with $m_t = 175\,\mathrm{GeV}/c^2$, $m_Z = 91.19\,\mathrm{GeV}/c^2$, $\Lambda = 1\,\mathrm{TeV}$ and $\alpha(M_Z) = (128.9)^{-1}$ and inserting the best branching ratio limits, following results hold:

$$\begin{aligned} \kappa^\gamma_{tq} &< 4.00 \cdot 10^{-2} \\ \kappa^Z_{tq} &< 0.102 \\ \sqrt{|v^Z_{tq}|^2 + |a^Z_{tq}|^2} &< 2.31 \cdot 10^{-2}, \end{aligned} \tag{4.13}$$

where q can be a u or c quark. The top width $\Gamma_t = 1.52\,\mathrm{GeV}$ here has been assumed.

Chapter 5

Conclusions

The top quark has been the last quark to be discovered and, though no discrepancies with the Standard Model have emerged yet, still many mysteries are hidden in this particle. Top quark is remarkably different from all the others, because it decays before any hadronization occurs: therefore, what is produced in its final state can say something fundamental, that would test the precision of Standard Model parameters or enlighten the occurrence of new physics. The fact that the huge top mass induces only small corrections to the three level diagram indicates that top quark is an ideal tool to search for something beyond the Standard Model.

Flavour Changing Neutral currents (FCNC) provide an example of these searches. The neutral couplings tVq with the $q = u, c$ quarks, where V is a gluon, a photon or a Z is not possible at tree level; at one-loop they are induced by charged-current interactions, which are GIM-suppressed. These contributions limit the FCNC decay branching ratios to extremely small values in the SM. Nevertheless, there are extensions of the SM which predict the presence of FCNC contributions already at the tree level and significantly enhance the top FCNC decay branching ratios.

The Large Hadron Collider offers the chance to observe such processes. Its high instantaneous luminosity, if combined with the large top pair and single top production at $\sqrt{s} = 14\,\text{TeV}$, determines a very high rate for the top production, thus allowing an extensive search for its different decaying mode. An observation of decays of the FCNC type would signal with any ambiguities that new physics is at work.

How well CMS – one of the two general purpose experiments at LHC – will be capable to attach this issue, is the object of the present work. In order to evaluate the CMS discovery potential for FCNC top decays, the $t \to Zq$ and $t \to \gamma q$ channels have been studied. Selection procedure has been optimized for a dataset corresponding to an integrated luminosity of $\text{L} = 10\,\text{fb}^{-1}$, using Monte Carlo data which has been processed using the full detector simulation.

A cut-based analysis has been defined, using objects reconstructed by the software packages developed by the CMS Collaboration. In particular, the selection procedure includes

an extensive set of quality requirements on the final state lepton candidates and is heavily reliant on good b-tagging capabilities. This procedure has been shown to be powerful in reducing the large background contributions from Standard Model $t\bar{t}$, QCD multi-jet and Z^0 production in association with a $b\bar{b}$ pair, while retaining a reasonable selection efficiency and minimizing the impact of systematic uncertainties.

The present study introduces for the first time some new features in the literature, thus resulting fairly original in several points. They are:

- most of the analysis is developed in the full simulation. When it had not been possible, a framework for fast simulation that has been fully validated against full simulation has been adopted. This is an important feature of the CMS reconstruction framework; namely, all considerations on experimental sensitivity to FCNC couplings drowned with the use of full simulation, are pretty fresh;

- a wide variety of Standard Model processes have been considered into background, detailing the impact of the selection for each of them. The impact of the same analysis on possible new physics channels, different from the one of concern, has been discussed;

- all the selection procedures have been devised in order to be poorly correlated each other, to maximize the significance region for the FCNC upper limit that has the best statistical properties, for both discovery limits and non-observation confidence level;

- all the selection procedures have been devised in order to be equally sensitive to $t \to uV$ and $t \to qV$ processes;

- all the selection procedures have been devised in order to be few susceptible to systematic effects. A full account of the impact of systematic effect (coming from detector imprecision and limited theoretical knowledge) has been studied, in a quite richer detail than in previous LHC studies;

- the analysis has considered to be suitable to extrapolate at integrated luminosity well higher than the standard $10\,\text{fb}^{-1}$. Results have extended to LHC in the high luminosity phase and the effect of the pile-up in the SLHC scenario has been addressed.

CMS has demonstrated to be capable to improve the current experimental limits for FCNC top decays by about two orders of magnitude, and by more than one order of magnitude even respect to the next-to-come results from Tevatron and HERA machines. The possibility to test these limits against a wide spectrum of theoretical models – that will hopefully be constrained by several other measures – opens up.

It is important to observe that such study can inspire analyses on other final states originated by the top FCNC. Top decays into neutral Higgs ($t \to cH$, $H = h^0, A^0, H^0$), which theory expects to have favourable branching ratios, could be inquired for some specific Higgs channels (as ZZ, WW, $\gamma\gamma$), adopting much of the apparatus deployed here. Even more

clearly, new physics could be searched in three-body decays as $t \to cWZ$, by applying few modifications to the $t \to qZ$ analysis. Putting constraints on new physics from these different manifestations is undoubtedly an exciting challenge.

As the bottom line of the work, it is worth to stress how studies of these effects are necessary. Due to the fact that different theories predict different orders of enhancement, the measurements of such processes at the LHC will not only shed light on new physics, but also may give some favour for a specified model. Once the nature of the virtual effects will be realized, it should be possible to infer if FCNCs are triggered by supersymmetry, alternative renormalizable extensions of the SM or even dynamic EWSB models, that in principle are not supposed to co-exist.

Signs of this new physics could be certainly searched for directly but, even if accessible, the corresponding signatures could be far from transparent. In contrast, the indirect approach based on the FCNC processes has the advantage to deal all the time with the dynamics of the top quark. Therefore, by looking for new features beyond the Standard Model properties of top quark, one could uncover the existence of new interactions. LHC is built to search the unexpected, and is waiting round the corner.

Acknowledgments

This work comes at the end of the author's PhD training. They have been years of exciting learning, hard studies, crazy rushes alternated to some dead time. The author would like to thank firstly Guido Tonelli, that involved him in CMS experiment, provided financial support for scientific formation, took care of work improvements and trigger efforts with his enthusiasm and team-leading skills. A special mention is needed to Andrea Giammanco, that trained several parts of the work, shared his wide experience and went through the whole text, being at disposal for any problem at any hour in the day or in the night. Acknowledgments are also for Roberto Tenchini and Paolo Azzurri that found some time for important suggestions, as well as Fabrizio Palla that gave the first guidelines for the study. This work would not have been possible without the efforts of Aris Kyriakis, that shared the bulk of analysis code with the author, succeeded in making the Athens and Pisa groups to meet together and issued a friendly and fruitful collaboration. The author is also indebted to Prof. Mauro Dell'Orso, Prof. Giacomo Polesello and Prof. Tommaso Tabarelli for their valuable comments in refereeing this thesis.

Author's sweetest thanks are for the immeasurable patience of his family, the attendance of his better half Delia Volpi, the help of few colleagues, the support of friends and all the souls that shared some feelings with him, in these years of sacrifices and passion for more and more learning.

Appendix A

Appendix

A.1 The role of m_t in EW precision physics

In Figure A.1 the updated results from the CDF Tevatron experiment are displayed. The final estimate is a (preliminary, Spring 07) combination of published Run-I results from the three decay modes with the same modes from Run-II, that benefits from a richer statistic and improved reconstruction techniques.

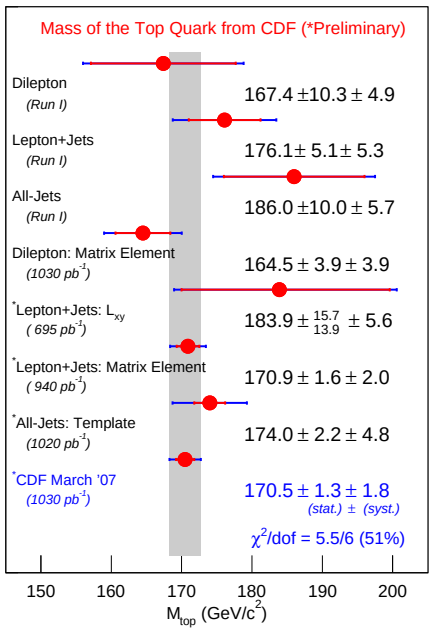

Figure A.1: CDF last result (Spring07) for the top mass. Combined with the similar one from DØ, it yields the estimate in Eq.1.1.

The fact that the top quark mass enters the EW precision observables as an input parameter via quantum effects (loop corrections) has some deep consequences. The large numerical

value of m_t gives rise to sizable corrections ($\Delta\rho$) that are proportional to m_t^2, *i.e.* [147] :

$$\Delta\rho(m_t^2) = N_C \frac{1}{16\pi} \frac{\alpha}{\sin^2\theta_W \cos^2\theta_W} \frac{m_t^2}{M_Z^2} - \frac{11}{24} \log\frac{m_H}{M_Z} \tan^2\theta_W,$$

where α is the fine structure constant, θ_W the weak mixing angle, N_C the number of QCD colours and M_Z the mass of the Z^0 boson. This is in contrast to the dependence on the mass of the Higgs boson (not explicited here), that is only logarithmic in leading order and therefore much weaker. This strong dependence of the SM radiative corrections on the input value of m_t is what made it possible to predict the value of m_t from the precision EW observables, before its actual experimental discovery.

The most important consequence of this sensitive dependence is that precision measurements of top quark mass, combined with W boson mass and exploiting theory relationships, provide a stringent test of the Standard Model.

The constraint on the mass of the Higgs boson is of particular interest. This parameter can be predicted from the top and W boson mass (M_W) via the following formula:

$$M_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F} \frac{1}{\sin\theta_W \sqrt{1-\Delta r}} \quad , \tag{A.1}$$

The quantity Δr is related to radiative corrections as $\Delta r = \Delta\alpha - \frac{\cos\theta_W^2}{\sin\theta_W^2}\Delta\rho$ (at one-loop level). $\Delta\rho$ contains the quadratic dependence on the top mass and the logaritmic dependence on the Higgs boson mass. The precision observables M_W and $\sin^2\theta$ are currently known with experimental accuracies of 0.05% and 0.07%, respectively [21]: exploiting these data, along with the current precision of the top mass and constraints from the Eq.A.1, stringent limits on the Higgs boson mass can be estabilished.

The figure A.2 shows the $\Delta\chi^2$ curve derived from a large amount of precision electro-weak measurements (performed at LEP, SLC, Tevatron), as a function of the Higgs boson mass, assuming the Standard Model. The preferred value (corresponding to the minimum of the curve) is $m_H = 73^{+20}_{-16}\,\mathrm{GeV}/c^2$ (at 68%, black line) and the theoretical uncertainty is indicated by the blue band into account. 'Low Q^2 data' rapresents the effect of W mass measurement from the NuTeV collaboration, that shows a $2.6 \div 2.8\sigma$ deviation from the other indirect constraints. The same precision electroweak measurements excludes a Standard-Model Higgs boson with mass greater than $144\,\mathrm{GeV}/c^2$ (one-sided 95% C.L.). This limit increases to $182\,\mathrm{GeV}/c^2$ when including the LEP 2 direct search limit ($114.4\,\mathrm{GeV}/c^2$, 95% C.L.) shown by the yellow exclusion area [148] . Therefore, the precision mass measurements of W boson and top quark allow the Standard-Model Higgs boson mass to be restricted to a small range of values.

Simulation of the ATLAS and CMS experiments have shown that, if a good control of systematic effects will be reached, a resolution about 25 MeV for M_W and $1 \div 2\,\mathrm{GeV}/c^2$ for m_t is at hand. Further improvements are expected from a combination of the two experiments. The precision on these two fundamental variables is going to become better and better, thus

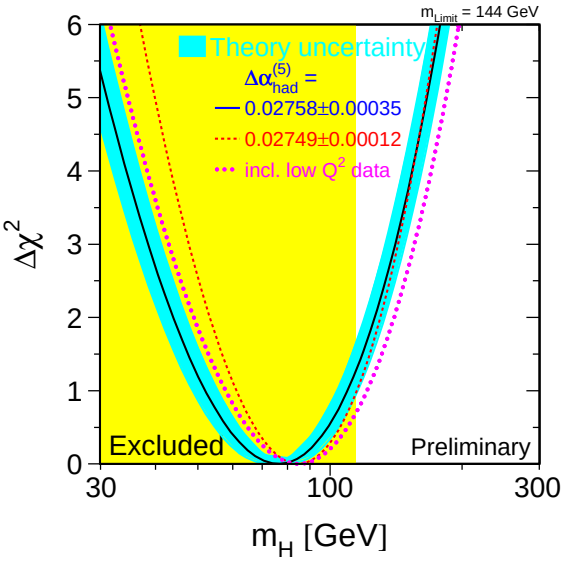

Figure A.2: The χ^2 value for the Higgs boson mass, including all direct and indirect measures performed up to now. Fluctuations come from theoretical and α_S uncertainties and inclusion of data from a low energy experiment, not well in agreement. Indirect EW data indicate the 68% confidence level for $m_H = 73^{+20}_{-16}\,\mathrm{GeV}/c^2$ and direct searches exclude the region where $m_H < 114.4$ GeV/c^2 .

restricting more and more the range of m_H and providing hints about the correct model of Nature.

A.2 Details about CMS sub-detectors

A.2.1 The solenoidal magnet

The required high momentum resolution for charged particles ($\Delta p/p \sim 0.1\ p_T$[TeV]) and the unambiguous determination of muon sign even for very large p_T, can only be assured by a long lever arm and a strong bending power.

The magnetic coil has been shaped trying to maximize the lever arm (that starts in the primary vertex, where the interesting particles tipically originate) and to ensure good momentum resolution even in the forward region. The high magnetic field has to be produced by a superconducting solenoid.

The main features of the CMS solenoid are the use of a high-purity aluminium-stabilized conductor and indirect cooling, together with full epoxy impregnation. The baseline technique was inherited from similar LEP and HERA experiment, but the unprecedent dimensions imposed several innovations. In particular, a four-layer winding has been adopted using a novel conductor with a larger cross section that can withstand an outward pressure (hoop

stress) of 64 atmospheres. The conductor carries a current of 20 kA and has a compound structure. The Rutherford-type cable is co-extruded with pure aluminium, which acts as a thermal stabiliser. This "insert" is then electron-beam-welded to 2 plates made of a high-strength aluminium alloy, for the mechanical reinforcement. The overall conductor cross section is 64×22 mm^2. The conductor was manufactured in twenty continuous lengths, each with a length of 2.65 km. Four lengths were wound to make each of the 5 coil modules, that were assembled and connected together.

With the field at the design value, the superconductong coil stores an energy of 2.7 GJ.

It provides a bending power of about 12 T·m in the barrel region $|\eta| < 1.45$, decreasing to about 4 T·m around $|\eta| \sim 2.4$.

A.2.2 The muon spectrometer system

The muon spectrometer is located within the iron yoke, where the return field can reach about 1.2 T and the bending power varies with η between 3 and 0.6 T·m.

The large bending power and the multiple scattering due to the amount of material inside the muon spectrometer lead to relative modest requirements on detector spatial resolution and alignement. A chamber resolution of the order of 100 μm is enough to improve the transverse momentum resolution for p_T >200 GeV/c, while the inner tracker dominates the resolution for lower momenta. For low-momentum muons, multiple scattering effect in the material before the first muon station dominates, and the best momentum resolution (by an order of magnitude) is given by measurements in the silicon tracker. Measurement of muons p_T using only the muon system is essentially determined by the muon bending angle at the exit of the coil, taking the primary interaction point (which will be known with a $\sim 20\,\mu$m precision) as the origin of the muon. The transverse momentum resolution using the muon system only, the inner tracker only and both ("full system") are displayed in Fig.A.3.

The CMS muon system [149] consists of three indipendent sub-systems, which structure and performance are briefly sketeched below. Each sub-system operates within the first level trigger system, providing 2 independent and complementary sources of information. The complete system results in a robust, precise and flexible trigger device.

Drift tube chambers

In the barrel region ($|\eta|$ <1.2), where the neutron induced background and muon rate is low and there is only a weak residual magnetic field, the drift tube (DT) chambers can be used.

DT cells have a section of 42×13 mm^2 and operate with an Ar/CO_2 mixture at atmospheric pressure. The maximum drift length is 2.0 cm and the single-point resolution is about 200 μm.

Four staggered layers of parallel cells ("superlayer") allow the left-right ambiguity of a single layer to be resolved (muon crosses at least three superlayers), provide the measurement

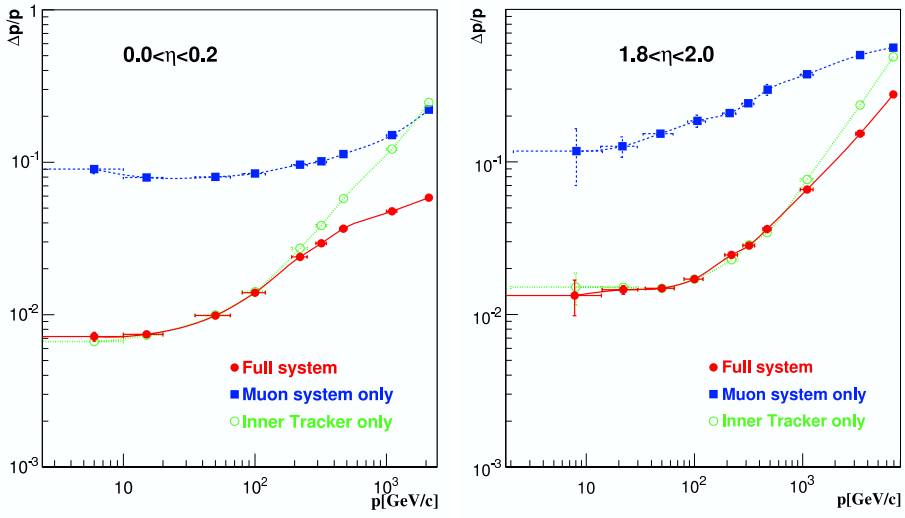

Figure A.3: The muon momentum resolution versus p_T using the muon system only, the inner tracker only and their combination, in the barrel (left) and endcap (right).

of a two-dimensional segment and measure the bunch crossing that has given origin to a segment, with no need of external input.

The Muon Barrel include 250 DT chambers, organized in 4 concentrical layers inside the magnet return yoke. The barrel is composed of 5 big rings, with each ring divided in 12 angular sectors ($\Delta\phi = 30°$). Muon detectors in each sector (called "stations") are built from two superlayers with wires in the z direction (to measure in $r\phi$ plane), most of which are interleaved with a superlayer with orthogonal wires (to measure in rz plane). Each station is designed to give a muon vector in space, with a resolution better than 100 μm in position and approximately 1 mrad in direction.

Cathode strip chambers

In the endcaps ($|\eta|$ <2.4), the muon and neutron background rate increase rapidly with pseudorapidity and the magnetic field is higher than in the barrel yoke. Cathode strip chambers (CSC) operating with an $Ar/CO_2/CF_4$ mixture at atmospheric pressure have a spatial and time resolution similar to DT, but they are capable to work even in some large inhomogeneous magnetic field and at high occupancy levels.

A CSC is composed of six trapezoidal layers of gas gap. Each gap has a plane of concentrical cathode strips (measuring the bending coordinate) and a plane of anode wires running almost perpendicularly to the strips (measuring the non-bending coordinate).

The gas ionization and subsequent avalanche caused by a charged particle traversing a chamber produces a charge on the anode wire and an image charge on a group of cathode strips, in each plane. The signal on the wires is fast and is used in the Level-1 Trigger, while the centre-of-gravity of the charge distribution induced on the cathode strips provides the most precise position measurement. Each CSC measures up to 6 (r, ϕ, z) coordinates, with a

spatial resolution between 100 and 200 μm and an angular resolution of order 10 mrad.

The whole system comprises 234 CSCs for each endcaps. CSC chambers are arranged in four disks (stations) placed between the iron disks of the yoke. Most of CSC are overlapped in ϕ, to avoid any gap in the muon acceptance. Data between DT and CSC are exchanged in the overlap region ($0.8< |\eta| <1.2$).

Resistive plate chambers

Redundancy is obtained with a system of Resistive Plate Chambers (RPC) that are installed both in the barrel and in the endcaps up to $|\eta| <2.1$. RPCs have limited spatial resolution but fast response and excellent time resolution ($\sim$2 ns), providing an unambigous bunch crossing identification. They are also used to complement the measurement of muon p_T, especially in the endcap where they can resolve spatial and temporal ambiguities in the CSCs.

These RPCs are operated in avalanche mode to ensure good operation at high rates (up to 10 kHz/cm^2) and have double gaps with a gap of 2 mm. They are rather less vulnerable to muon radiation than both DT system (that have a quite long drift time, about 400 ns) and CSC (that suffer from charge weighting).

In the barrel, each DT chamber has 1 or 2 RPCs coupled to it before installation. In the endcap, each of two rings of each station features 36 chambers.

A.2.3 The CMS calorimetric system

The CMS calorimetric system is located inside the 4 T coil of the solenoid magnet. It is made by the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL) and ensures an energy resolution better than 1% for electrons and photons around 100 GeV.

The ECAL consists of about 76,000 $PbWO_4$ scintillating crystals, arranged in a barrel part (EB) in the $|\eta| <1.48$ region and two endcap (EE), which extends the coverage up to $|\eta| <3$.

The HCAL is a sampling device divided into four kind of hadronic calorimeters, which provide good segmentation and hermeticity, moderate energy resolution and full angular coverage up to $|\eta| <5$:

- the Barrel HCAL (HB) surrounds the EB, covering the central pseudorapidity region up to $|\eta| <1.3$ with a 5.8 λ_I depth. The HB modules consists of 17 layers of plastic scintillator interleaved with brass absorber. Layers are segmented into $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ towers and provide one depth measurement;
- Outer HCAL (HO) consists of one/two layers scintillators located outside the magnet, that extend the central shower containement to a 11.8 λ_I depth;
- the Endcap HCALs (HE) cover the region up to $|\eta| < 3$ and have a depth segmentation varying from one to three;

- two forward calorimeters (HF) surround the beam pipe 11 m from the interaction point. They are made by quartz fiber to operate in this very hard radiation area and extend coverage up to $|\eta| = 5$.

A.2.4 CMS Electromagnetic calorimeter

The Electromagnetic Calorimeter (ECAL, [150]) is a hermetic, homogeneous calorimeter comprising 61,200 $PbWO_4$) crystals mounted in the central barrel part (ECAL barrel, EB), closed by 7324 crystals of the same material in each of the 2 endcaps (ECAL endcap, EE).

The use of lead tungstate crystals (featuring a very short radiation lenghth X_0 =0.89 cm and Molière radius R_M =2.2 cm) allows the construction of a very compact, higly granular and radiation resistant (up to 10 MRad) detector. In addition fast time response (80% of the light is emitted within 25 ns) is assured.

Neverthless, the relatively low light yield (∼30 photons/MeV) imposes the use of photodetectors with intrinsic amplification and capable to operate in a so strong magnetic field. Silicon avalanche photodiodes (APDs) are used in the barrel and vacuum phototriodes (VPTs) are preferred in endcaps due to higher neutron flux. The sensitivity of both the crystals and the APD response to temperature changes requires a temperature stability of 0.1°C at least.

Each crystals is wedge-shaped, with a 25.8 X_0 lenghth in EB (24.7X_0 in EE) and a front face cross section of 2.2×2.2 (2.9×2.9) cm^2. Photodetectors are glued to the rear face.

In the EB ($|\eta|$ <1.479), crystal granularity amounts to $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$; to avoid gaps in acceptance they are not exactly projective to the nominal vertex position (axes tilted by 3°). The whole barrel section has an inner radius of 129 cm and is structured as 36 identical "supermodules", each covering half the barrel length and including 1700 crystals.

In the EE (1.479< $|\eta|$ <3), crystals are arranged in an x-y grid, similarly off-pointing from the nominal vertex position and with the same granularity. Each endcaps is partioned in two "Dees" (semicircular plates), 314 cm distant from the nominal vertex.

A "preshower" device is placed in front of the crystal calorimeter over much of the endcap pseudorapidity coverage. The active elements of the Preshower are 2 planes of silicon strip detectors, with a pitch of 1.9 mm, which lie behind disks of lead absorber at depths of $2X_0$ and $3X_0$. The energy measured by the preshower has to be added to the crystals supercluster energy.

All crystals are grouped in 5×5 matrix called supercrystals. Precision energy measurements can be performed up to $|\eta|$ <2.6, where the ECAL signals can be effectively matched with the tracker ones.

Performance of the electromagnetic calorimeter

The performance of several supermodules were measured in some beam tests with electron of defined energy. The energy resolution σ/E (where σ is that from a Gaussian function fitting

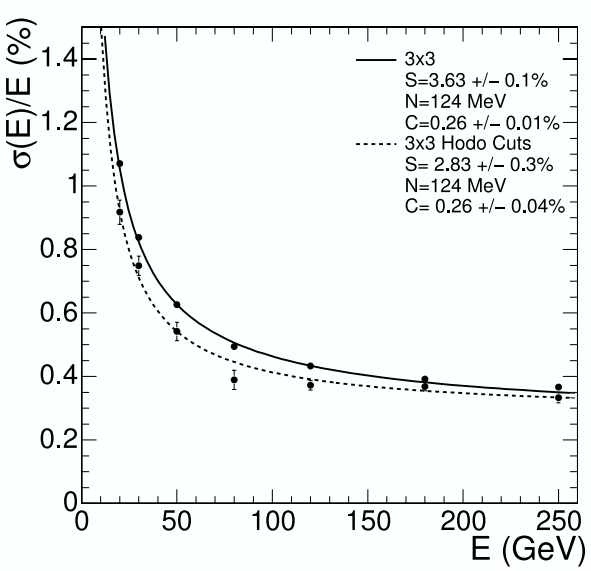

Figure A.4: ECAL supermodule energy resolution, as a function of electron energy, as measured from a beam test. The energy was measured in an array of 3×3 crystals with electrons impacting the central crystal. The upper series of points correspond to events taken with a trigger in a $20\times20\,\mathrm{mm}^2$, while the lower series are requested to fall within a $4\times4\,\mathrm{mm}^2$ region.

the reconstructed energy distributions) is expected to run with E following the expression:

$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C, \tag{A.2}$$

i.e. as the quadratic sum of a stocastic, noise and constant terms. One of the result of the test is shown in Fig. A.4. It exhibits a nice match with the function. Fitted parameter value are displayed inside the picture.

A.2.5 Hadron calorimeter

The design of the hadron calorimeter (HCAL, [151]) is strongly influenced by the fact that most of absorbing material having to be maximized (in terms of interaction lengths) is inside the magnet coil. Important requirements to the HCAL are to minimize the non-Gaussian tails in the energy resolution and to provide good containment and hermeticity for the missing energy detection.

Brass has been chosen as absorber material in most of HCAL part, as it has a reasonably short interaction length, it is easy to machine, widely available and non-magnetic.

Maximizing the amount of absorber before the magnet requires keeping to a minimum the amount of space devoted to the active medium: therefore, tile/fibre technology seems an ideal choice. It consists of plastic scintillator tiles with a 3.7 mm thickness, red out with embedded wavelength-shifting fibres. The photodetection readout is based on multi-channel hybrid photodiodes (HPDs).

Brass plates have been precisely bolted together before insertion, so no un-instrumented cracks or dead areas in ϕ are expected. The gap between the barrel and the endcap HCAL (through which the services of the ECAL and the inner tracker pass) is inclined in a way that avoids to point to the interaction vertex.

Hadron barrel

The barrel part ($|\eta|$ <1.4) of the detector (HB) surrounds the EB system and is attached to the inner wall of the vacuum vessel of the coil. It was assembled in two half barrels, each partitioned in 18 wedges whose individual weight is around 28 tonnes. The absorber consists of a 40 mm thick front steel plate, followed by eight 50.5 mm thick brass plates, six 56.5 mm thick brass plates, and a 75 mm thick steel back plate. This result in 2304 towers with a segmentation $\Delta\phi \times \Delta\eta = 0.087 \times 0.087$. No longitudinal segmentation is done, and the HB is red as a single longitudinal sampling.

Hadron outer

The hadron barrel is complemented by an additional layer of scintillators lining the outside of the vacuum tank of the coil, referred to as the hadron outer (HO) detector.

HO samples the energy from penetrating hadron showers leaking through the rear of the calorimeters, serving as a "tail-catcher" of the energy resolution function. As a result, the effective thickness of the hadron calorimetry is increased to over 10 interaction lengths. Also E_T^{miss} resolution of the calorimeter is improved.

Scintillators in this device have a 10 mm thickness and match the ϕ segmentation of the DT chambers. They cover the $|\eta|$ <1.26 region.

HO is physically located inside the barrel muon system and is hence constrained by its geometry and construction of that system. It is divided into 5 "rings" along η, with the central one having 2 scintillator layers and the other a single layer, intespersed in the iron absorber.

Hadron endcap

The 1.3< $|\eta|$ <3.0 pseudorapidity region is covered by the hadron endcap (HE), where the same technology of HB is implemented. A total of 14 towers for endcap is fixed on the inner yoke disk, pointing to the interaction vertex. Granularity ranges from $\Delta\phi\times\Delta\eta = 0.087\times0.087$ in the outermost towers to $\Delta\phi \times \Delta\eta = 0.174 \times 0.035$ in the innermost ones.

Hadron forward

Coverage in the pseudorapidities of 3.0< $|\eta|$ <5.0 is provided by the steel/quartz fibre Hadron Forward (HF) calorimeter.

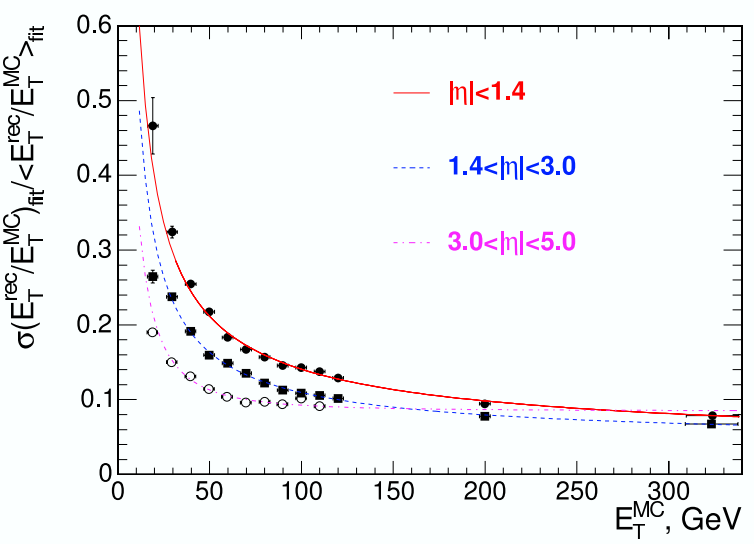

Figure A.5: The jet transverse energy resolution as a function of the simulated jet transverse energy E_T^{MC} for barrel jets ($|\eta| < 1.4$), endcap jets ($1.4 < |\eta| < 3.0$) and forward jets ($3.0 < |\eta| < 5.0$). Comparison is done with the reconsructed jet energy E_T^{rec}.

The congested hadron rate in this region is managed by the two absorbers, located at 11.2 m from the interaction point. The HF technology samples preferentially the neutral component of the hadron shower, that leads to narrower and shorter jets: thus the depth of the absorber can be limited to 1.65 m.

The signal originates from Cerenkov light emitted in the quartz fibres, which is then channelled by the fibres to photomultipliers. The quarz fibers have a 0.6 mm diameter and they run parallel to the beam line (with two different lengths creating 2 longitudinal samplings). The absorber structure is created by machining 1 mm square grooves into steel plates. Segmentation ranges approximatively from $\Delta\phi \times \Delta\eta = 0.174 \times 0.175$ to $\Delta\phi \times \Delta\eta = 0.35 \times 0.3$, for a total amount of 900 towers.

Performance of the hadron calorimeter

The granularity of the sampling in the 3 parts of the HCAL has been chosen such that the jet energy resolution, as a function of E_T, is similar in all 3 parts, as illustrated in Figure A.5. All the jets are reconstructed with the 'iterative cone' $R = 0.5$ algorithm, that will be explained later.

A.2.6 Inner tracking system

By considering the charged particle flux at various radii at high luminosity, 3 regions can be identified:

- few centimeters in radius around the interaction vertex, where the particle flux is the

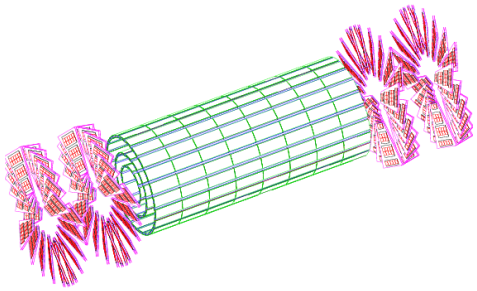

Figure A.6: Layout of pixel detectors in the CMS tracker.

highest (order of 10^7/s). Here a pixel detector with pixel of sub-millimeter size assures an occupancy below 10^{-3};

- intermediate region (about 20< r <60 cm), where the particle flux is low enough to enable use of silicon microstrip detectors, with a pitch around 100 μm and a length around 10 cm, leading to an occupancy of 2-3%;

- an outermost region where the particle flux has dropped sufficiently. Silicon strips are still needed, but now relaxed requirements on the pitch and the length are adequate for a 1% occupancy.

Putting these choices together has resulted in filling the inner CMS volume with entirely silicon-based detector, covering the 4< r <120 cm radial and $|\eta|$ <2.4 pseudorapidity region. The total length is approximately 540 cm and silicon sensor cover a 1 m^2 surface in the pixel and 200 m^2 in the strip region. The inner tracker comprises 66 million pixels and 9.6 million strips.

Pixel tracker

The Pixel detector grants the most accurate spatial measurements in the Tracking system, providing a three-dimensional position information. In addition, it is caracterized by a very low occupancy (at most $\mathcal{O}(10^{-4})$ hits per pixel at each bunch crossing, at the design luminosity) even in the high density environment of p-p LHC collision.

In order to achieve the optimal vertex position resolution in both (r, ϕ) and z, pixel elements have a size 100x150 μm^2. The detector is readout using approximately 16,000 readout chips, which are bump-bonded to the detector modules.

The pixel detector element are arranged in 3 layers in the barrel and 2 disks in the forward parts (Fig. A.6) for a total about 66x10^6 readout channels.

The barrel layers are located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm and have a length of 53 cm. The 2 end disks (extending from 6 to 15 cm in radius) are placed on each

side at $z = \pm 34.5$ cm and $z = \pm 46.5$ cm. The barrel comprises 768 pixel modules arranged into half-ladders of 4 identical modules each. The large Lorentz effect (Lorentz angle is 23°) improves the $r - \phi$ resolution through charge sharing. The endcap disks are assembled in a 'turbine-like' geometry with blades rotated by 20° to also benefit from the Lorentz effect. The endcap disks comprise 672 pixel modules (called "plaquette") with 7 different modules in each blade.

Spatial resolution is about 10 μm in the (r, ϕ) plane and about 20 μm in the z direction, but it strongly depends on the track impact angle and the size of clusters.

The Silicon Strip Tracker

The barrel tracker region is divided into 2 parts: a TIB (Tracker Inner Barrel) and a TOB (Tracker Outer Barrel).

The TIB is made of 4 layers and covers up to $|z|$ <65 cm, using silicon sensors with a strip pitch varying from 80 to 120 μm. The third and fourth layers are made of single-sided module with a thickness of 320 μm, while in the first 2 layers single-sided sensors are glued back-to-back with a stereo tilt angle (100 mrad), in order to obtain a three dimensonal position measurements. This leads to a single-point resolution of between 23 and 34 μm in the $r\phi$ direction and 230 μm in z.

The TOB has 6 layers with a half-length of $|z| < 110$ cm. As the radiation levels are smaller in this region, thicker silicon sensors (500 μm) can be used to maintain a good S/N ratio for longer strip length and wider pitch. The strip pitch varies from 100 to 180 μm. Also for the TOB the first 2 layers provide a 'stereo' measurement in both $r\phi$ and rz coordinates. The single-point resolution varies from 35 to 52 μm in the $r\phi$ direction and 530 μm in z.

The endcaps are divided into the TEC (Tracker End Cap) and TID (Tracker Inner Disks). Each TEC comprises 9 disks that extend into the region 120 cm< $|z|$ <280 cm, and each TID comprises 3 small disks that fill the gap between the TIB and the TEC. The TEC and TID modules are arranged in rings, centered on the beam line, and have strips that point towards the beam line, therefore with a variable pitch ($80 \div 120$ μm in TID and $100 \div 220$ μm). The first 2 rings of the TID and the innermost 2 rings and the fifth ring of the TEC have 'stereo' modules. The thickness of the sensors is 320 μm for the TID and the 3 innermost rings of the TEC and 500 μm for the rest of the TEC.

The entire silicon strip detector consists of almost 15,400 modules [152], which will be mounted on carbon-fibre structures and housed inside a temperature controlled outer support tube. The operating temperature will be around $-15°$ C.

The schematic view in Fig. A.7 shows the position of silicon modules in the rz plane for each system (Pixel, TIB, TID, TOB, TEC).

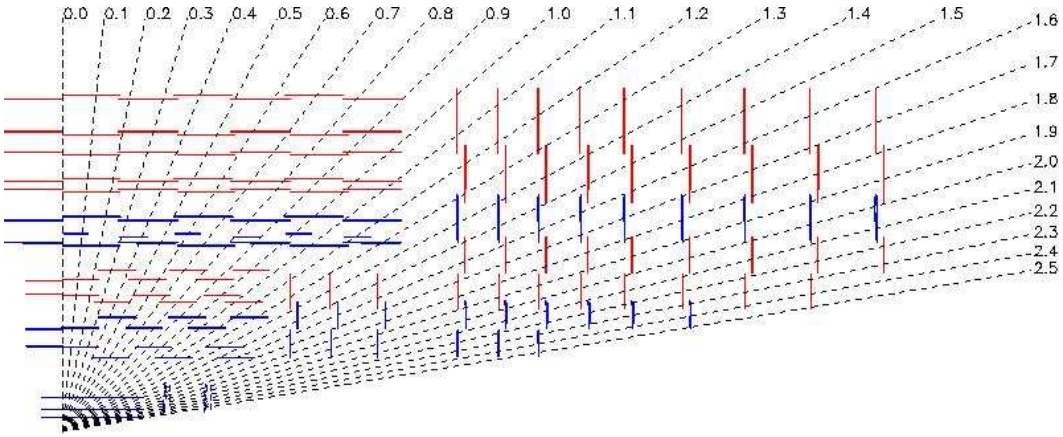

Figure A.7: Layout of a quarter of CMS tracker, with lines representing modules sides. With increasing r: Pixel, TIB, TOB. With increasing z: Pixel, TID, TEC.

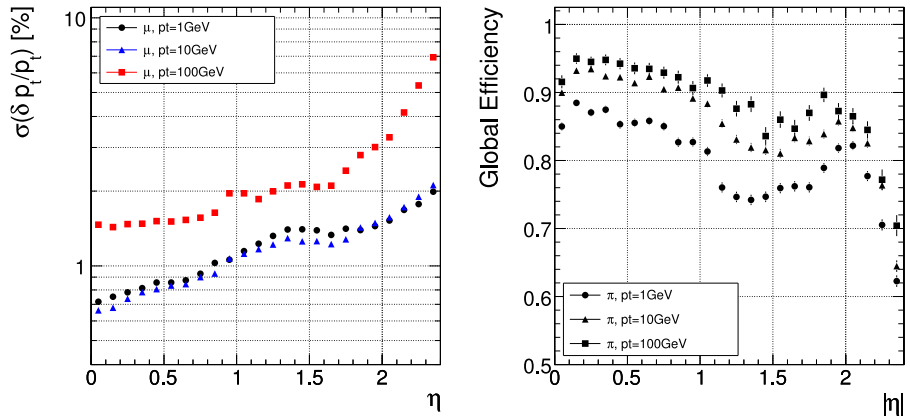

Figure A.8: Left: The relative p_T resolution of single muon as a function of η, for three reference energies. Right: the global reconstruction efficiency along pseudorapidity, for the same reference energies.

Track and Vertex reconstruction performance

In order to identify a b quark among the partons in the final state, a very precise track and vertex reconstruction is needed. The efficiency to reconstruct tracks depends on many factors, such as the event topology, detector efficiency and luminosity conditions. Single muon tracks are reconstructed with an efficiency close to 100% in the tracker acceptance, assuming a perfectly aligned detector. Figure A.8 (right) shows that track reconstruction efficiency for pions with the CMS detector simulation does not fall below 85%. The relative p_T resolution for single muon tracks is shown in Fig. A.8 (left) as a function of pseudorapidity.

Target resolution of few percent at high energy is sucessfully reached. The impact parameter resolution (evaluated on high p_T tracks) amounts to $\sigma(d_0)$ =20 μm for the transverse and $\sigma(z_0)$ =40 μm for the longitudinal direction.

All the results given above are obtained with algoritms based on a standard Kalman Filter method [107].

The vertex finding process is accomplished in two steps: firstly primary vertex are reconstructed, identifying the tracks directly coming from it (usually those that triggered the event). Subsequently the reconstruction of displaced vertices from high lifetime b quark is performed. The performance on vertex reconstruction is clearly related to the quality of the track reconstruction.

The reconstruction of primary vertices can also be performed at an early stage, without using the whole information from all the tracking system but only the Pixel detector. The main advantage of such an approach is that it is fast, so it can be used for event selection at trigger level and to costrain the track reconstruction from the full tracker.

The resolution in the z position determination is about $50\,\mu$m for the low luminosity phase and improves to $30\,\mu$m using the full tracker information.

To reconstruct secondary vertices, the informations from all tracker sub-systemsle are needed. The efficiency of secondary vertex finding depends on the impact parameter of tracks belonging to a displaced vertex and on the purity on the same set of tracks. Once the same set of tracks coming from a displaced vertex have been identified, a fit is needed to estimate the position of the secondary vertex from which the decay length is computed.

A.2.7 CMS forward detectors

As outlined in Sec. 2.2, central detectors of CMS experiment have an acceptance in pseudorapidity of roughly $|\eta|$ <2.5 for tracking information and $|\eta| < 5$ for calorimeter information. If one would catch the most part of the energy in the collision this is still not sufficient, because the greater fraction of charged particles and of the energy flow at the LHC is produced close to the beam pipe (about 5< $|\eta|$ <11). Hence presently there are 2 proposals to extend the coverage in the forward region, that will be briefly presented below.

The TOTEM experiment

The TOTEM experiment [153] has been conceived to measure the pp elastic cross section as a function of the exchanged four-momentum, the total cross section (with a precision of approximately 1%) and diffractive dissociation at $\sqrt{s} = 14$ TeV.

The TOTEM experimental set-up (sketeched in Fig. A.9) consists of 2 tracking telescopes T1 and T2 along with a Roman Pot (RP) stations, one on either side of IP5. The T1 and T2 telescopes are made of Cathode Strip Chambers and GEM (Gas Electron Multipliers) chambers respectively, and will detect charged particles in the regions 3.2< $|\eta|$ <5 and 5< $|\eta|$ <6.6. The TOTEM RP stations will be placed at a distance of ±147m and ±220m from IP5.

The TOTEM detectors can provide input data to the Global Trigger of the CMS Level-1 trigger. Track finding in T1 and T2 for triggering purposes is optimized to select beam-beam events with charged particle tracks that point back to the IP, thus rejecting beam-gas and

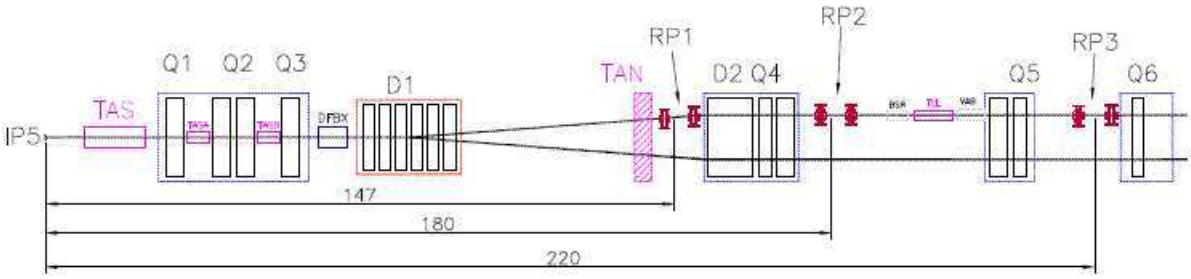

Figure A.9: The LHC beamline and the Roman Pots of TOTEM at 3 different locations. Distances are measured in meters. The detectors at 180 m are optional.

beam-halo events which have tracks that do not.

CASTOR and ZDC systems

CASTOR and the Zero Degree Calorimeter (ZDC) are calorimeters with both an electromagnetic and hadronic section, that has been proposed to cover the $5.1 < |\eta| < 6.5$ region. These calorimeters are of interest for measurements in pp, pA and AA collisions, where A is a heavy ion.

Both apparata are Cerenkov-light devices, consisting of successive layers of tungsten plates as absorber and fused silica (quartz) plates (fibers in ZDC) as active medium. The use of tungsten provides sufficient density to contain the transverse spread of the signal to within a few cm of the initial trajectory, while quartz scintillators assure high radiation tolerance. Cerenkov calorimeters are almost immune to induced radiation since they are only sensitive to high velocity charged particles.

CASTOR has a total depth about $22X_0$ in the EM section and 10.3 λ_I in total. It is situated in the collar shielding at the very forward region of CMS, starting at 14.37 m from the interaction point as shown in Fig. A.10.

The energy resolution function follows the standard behaviour of Eq. A.2, with a stocasthic term about $26\%/\sqrt{E[\mathrm{GeV}]}$ and constant term around 2.5%.

For the ZDC (with a 8 λ_I depth) beam tests on prototype show a resolution of 10% for 2.7 TeV neutrons and 11% for 50 GeV photons.

Thanks to the ZDC capability to precisely sample the narrow electromagnetic showers

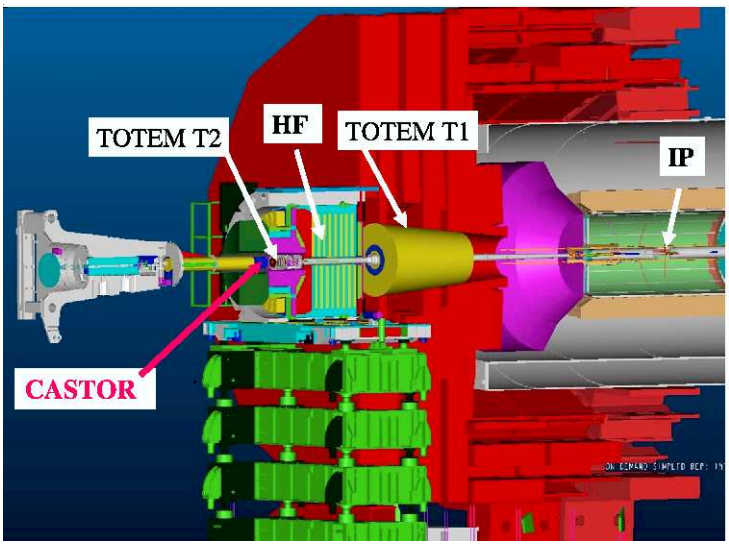

Figure A.10: Schematics of the CMS forward region.

(position resolution ≈6 mm) the beam crossing angle can be measured. Furthermore, a timing resolution of order 100 ps can be exploited to make a quick vertex selection (3 cm resolution) already at L1 trigger.

A.3 The SLHC machine and the physics programme

SLHC is an upgraded version of LHC, that is planned to run at an instantaneous luminosity of 10^{35} cm^{-2}s^{-1}.

Assuming that the physics programme for LHC will have been accomplished succesfully – in particular that the Higgs boson and supersymmetry will have been found in the mass ranges expected today – the upgraded SLHC would offer much more discovery potential. Here below some suggestions are summarised [98]:

- the measurement of some of the Triple Gauge Coupling will reach an accuracy comparable with the size of the electroweak (and possibly SUSY) virtual corrections;

- new rare decay modes of the SM Higgs boson will become accessible, as $H \to \mu^+\mu^-$ and $H \to Z\gamma$. The determination of the Higgs couplings to bottom and top quarks will reach precisions better than 10% over a large fraction of the Higgs boson mass range. In the MSSM, the region of SUSY parameter space where at least two Higgs bosons will be observed is significantly enlarged with respect to the LHC reach;

- the first observation of SM Higgs pair production may be possible in the 170< m_H <200 GeV/c^2 mass range, with a determination of the Higgs self-coupling λ_{HHH} at a level of 19% (25%) for $m_H = 170$ GeV/c^2 ($m_H = 200$ GeV/c^2), after background subtraction;

- in the absence of a Higgs signal, studies of resonant and non-resonant scattering of electroweak vector boson pairs at high mass will benefit from the larger statistics, which should give access to a larger variety of channels and in general to more convincing signals than at the LHC;

- the mass reach for squarks and gluinos will be extended to about 3 TeV and some exclusive SUSY channels, rate-limited at the standard LHC, could be studied in detail with a ten-fold increase in statistics;

- the mass reach for new gauge bosons, or for signatures of extra-dimension models, will be extended by about 30% relative to the LHC.

Finally, the possibility to study in detail all rare decays and production process in the top quark domain is fully at hand.

The SLHC challenging environment

The main drawback of a such high luminosity machine will be the very high radiation environment. All physics analyses could be hardly affected by the large event pile-up, have to cope with a reduced efficiencies and a possibly increased backgrounds. On another side, the detectors will have to deal with about 200 collisions per 25 ns, producing about 1200 charged particle tracks per unit of pseudo-rapidity.

The assumptions that today are adopted for the physics studies are outlined below:

Tracking. It has been assumed that, provided that a large part of the inner detectors of both experiments have to be replaced with more radiation hard and granular devices, reconstruction of isolated hard particles (as the muons and electrons involved in the analysis) will be possible with efficiency and momentum resolution comparable with the present detectors;

b-tagging. It has been assumed that with the new pixel detectors – that has to replace the original ones to cope with the enhanced radiation environment – the probability of confusion in the pattern recognition remains low, and the extra (fake) b-tags are given by real tracks from the minimum bias events, which are produced near the main event primary vertex and within the jet cone. The expectations for the mistagging with the u quark, for instance, range from 3.7% when b-jet transverse momentum is between 60 and 100 GeV/c and 0.88% when it is between 100 and 200 GeV/c: therefore, it lies inside what has been optimized here for the jet selection;

Electron identification and measurement. It has been calculated that an increase of a factor 10 in luminosity increases the contribution of the pile-up noise to the calorimeter energy resolution by about a factor 3. This deterioration is expected to be smaller for

electron energies of some tens of GeV, because the contribution of the pile-up noise to the energy resolution decreases with the particle energy as $1/E$;

Muon identification and measurement. If enough shielding can be installed in the forward regions to protect the muon spectrometers from the increased radiation background, the muon reconstruction efficiency and momentum resolution provided by the muon chambers are not expected to be seriously deteriorated when running at the SLHC;

Jet reconstruction. The increased pile-up can give rise to additional jets in the detector, that may spoil the efficiency of the jet reconstruction. The signal purity and background rejection could be recovered to some extent by increasing the jet thresholds or narrowing the cone sizes below $\Delta R = 0.2$. All the jet cones exploited in the present analysis satisfy this definition;

Trigger. The optimal efficiency for the inclusive triggers adopted in the study is expected to be almost fully recovered, with some increasing of the thresholds that are presently under discussion.

Some of these performance will be verified later, through the usage of the fast simulation framework. In all, the degradation when scaling from LHC to SLHC are not expected to be dramatic, thus the extrapolation of the present study seems affordable.

Bibliography

[1] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **73**, 225 (1994), hep-ex/9405005; F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **74**, 2626 (1995), hep-ex/9503002; S. Abachi *et al.* [DØ Collaboration], Phys. Rev. Lett. **74**, 2632 (1995), hep-ex/9503003; F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **51**, 4623 (1995), hep-ex/9412009; F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **75**, 3997 (1995), hep-ex/9506006; F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **52**, 2605 (1995).

[2] S. Abachi *et al.* [DØ Collaboration], Phys. Rev. Lett. **79**, 1197 (1997), hep-ex/9703008; F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **80**, 2767 (1998), hep-ex/9801014; B. Abbott *et al.* [DØ Collaboration], Phys. Rev. Lett. **80**, 2063 (1998), hep-ex/9706014; F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **80**, 2779 (1998), hep-ex/9802017; B. Abbott *et al.* [DØ Collaboration], Phys. Rev. D **60**, 052001 (1999), hep-ex/9808029; F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **82**, 271 (1999), hep-ex/9810029.

[3] [CDF Collaboration], Phys. Rev. Lett. **96**, 152002 (2006), hep-ex/0512070; [CDF Collaboration], Phys. Rev. D **74**, 032009 (2006), hep-ex/0605118; [CDF Collaboration], Phys. Rev. Lett. **96**, 022004 (2006), hep-ex/0510049; [CDF Collaboration], Phys. Rev. D. **73**, 032003 (2006), hep-ex/0510048; [CDF Collaboration], Phys. Rev. D. **73**, 092002 (2006), hep-ex/0512009; [CDF Collaboration], Phys. Rev. D. **75**, 111103 (2007); [DØ Collaboration], Phys. Rev. D **74**, 092005 (2006), hep-ex/0609053; [DØ Collaboration], Phys. Rev. D **75**, 092001 (2007); [DØ Collaboration], FERMILAB-PUB-07/039-E, Submitted to Phys. Rev. D February 9, 2007, hep-ex/0702018; [CDF and DØ Collaboration], hep-ex/0703034.

[4] F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **59**, 092001 (1999).

[5] S. Abachi *et al.* [DØ Collaboration], Phys. Rev. Lett. **79**, 1203 (1997), hep-ex/9704015; F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **80**, 2773 (1998), hep-ex/9710008.

[6] [CDF Collaboration], Phys. Rev. Lett. **93**, 142001 (2004); [DØ Collaboration], Phys. Lett. B **626**, 35 (2005), hep-ex/0504058; [CDF Collaboration], Phys. Rev. D **71**, 052003 (2005); [DØ Collaboration], Phys. Lett. B **626**, 55 (2005), hep-ex/0505082; [CDF Collaboration], Phys. Rev. D. Phys. Rev. D **75**, 111103 (2007), hep-ex/0607095; [DØ Col-

laboration], Phys. Rev. D **74**, 112004 (2006), hep-ex/0611002; [CDF Collaboration], Phys. Rev. D **75**, 111103 (2007); [DØ Collaboration], Phys. Rev. D **74**, 112004 (2006).

[7] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **79**, 1992 (1997); B. Abbott *et al.* [DØ Collaboration], Phys. Rev. D **60**, 012001 (1999), hep-ex/9808034.

[8] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **79**, 3585 (1997), hep-ex/9704007.

[9] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **84**, 216 (2000), hep-ex/9909042.

[10] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **71**, 031101(R) (2005); [DØ Collaboration], Phys. Lett. B 617, 1 (2005); [CDF Collaboration], Phys. Rev. D **73**, 111103(R) (2006) [DØ Collaboration], Phys. Rev. D Rap. Comm. **75**, 031102(R), (2007), hep-ex/0609045.

[11] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **95**, 102003 (2005) [DØ Collaboration], Phys. Lett. B **639**, 616 (2006), hep-ex/0603002.

[12] B. Abbott *et al.* [DØ Collaboration], Phys. Rev. Lett. **85**, 256 (2000).

[13] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **80**, 2525 (1998). V. Abazov *et al.* [DØ Collaboration], hep-ex/0702005, submitted to Phys. Rev. Lett. on 1 February 2007.

[14] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **79**, 357 (1999); [DØ Collaboration], Phys. Rev. Lett. **82**, 4975 (1999); [DØ Collaboration], Phys. Rev. Lett. **88**, 151803 (2002).

[15] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. Lett. **84**, 835 (2000) [DØ Collaboration], Phys. Rev. Lett. **92**, 221801 (2004). [CDF Collaboration], FERMILAB-PUB-0= 7-455-E. Submitted to Phys. Rev. Lett. September 6, 2007; [DØ Collaboration], DØ conference note 5443 (2007).

[16] [DØ Collaboration], Phys. Rev. D Rap. Comm. **63**, 031101(R) (2000); [DØ Collaboration], Phys. Lett. B **517**, 282 (2001); [CDF Collaboration], Phys. Rev. D **68**, 052003 (2004); [CDF Collaboration], Phys. Rev. D, **71** 012005(R) (2005); [DØ Collaboration], Phys. Lett. B **622**, 265 (2005), hep-ex/0505063.

[17] [DØ Collaboration], Phys. Rev. Lett., **98**, 181802 (2007), hep-ex/0612052.

[18] [CDF Collaboration], CDF Note 8964.

Phys. Rev. Lett., **98**, 181802 (2007), hep-ex/0612052.

[19] [ZEUS Collaboration], Phys. Lett. B **559**, 153 (2003).

[20] Phys. Rev. Lett. **98**, 041801 (2007), hep-ex/0608044 .

[21] The LEP Electroweak Working Group and the SLD Heavy Flavour Group, CERN-EP/99-15; D. G. Charlton, Int. J. Mod. Phys. A **1561**, 352 (2000), hep-ex/9912019; M. L. Swartz, Int. J. Mod. Phys. A **1561**, 307 (2000), hep-ex/9912026.

[22] M. Beneke *et al.*, (2000),hep-ph/0003033.

[23] Tevatron Electroweak Working Group (for the CDF and DØ Collaborations), hep-ex/0703034 (Submitted on 19 Mar 2007).

[24] W.-M. Yao *et al.* [Particle Data Group], Journ. of Phys. G **33**, 1 (2006), http://pdg.lbl.gov.

[25] G. Altarelli, L. Conti and V. Lubicz, Phys. Lett. B **502**, 125 (2001), hep-ph/0010090.

[26] R. Decker, M. Nowakowski and A. Pilaftsis, Mod. Phys. Lett. A **6**, 3491 (1991); G. Mahlon and S. Parke, Phys. Lett. B **347**, 394 (1995), hep-ph/9412250.

[27] P. Nason, S. Dawson, R. K. Ellis, Nucl. Phys. B **303**, 607 (1988) W. Beenakker, H. Kuijf, W. L. van Neerven and J. Smith, Phys. Rev. D **40**, 54 (1989).

[28] G. Sterman, Nucl. Phys. B **281**, 310 (1987); S. Catani and L. Trentadue, Nucl. Phys. B **327**, 323 (1989); S. Catani, M. L. Mangano, P. Nason and L. Trentadue, Nucl. Phys. B **478**,273 (1996), hep-ph/9604351, E.L. Berger and H. Contopanagos, Phys. Rev. D **57**, 253 (1998), hep-ph/9706206.

[29] N. Kidonakis and G. Sterman, Nucl. Phys. B **505**, 321 (1997), hep-ph/9705234; R. Bonciani, S. Catani, M.L. Mangano and P. Nason, Nucl. Phys. B **529**, 424 (1998), hep-ph/9801375; D. Chakraborty, J. Konigsberg, D. Rainwater, Ann. Rev. Nucl. Part. Sci. **53**, 301 (2003); N. Kidonakis, R. Vogt, Phys. Rev. D **68**, 114014 (2003).

[30] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan and S. Weinzierl, Phys. Rev. D **66**, 054024 (2002), hep-ph/0207055; Q. H. Cao and C. P. Yuan, Phys. Rev. D **70**, 094012 (2004), hep-ph/0408158; Phys. Rev. D **71**, 054022 (2005), hep-ph/0408180; J. Campbell, R. K. Ellis and F. Tramontano, Phys. Rev. D **71**, 054023 (2005), hep-ph/0409040; Q. H. Cao, R. Schwienhorst, J. A. Benitez, R. Brock and C. P. Yuan, Phys. Rev. D **72**, 094027 (2005), hep-ph/0504230; S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, JHEP 06 03, 092(2006), hep-ph/0512250.

[31] S.S. Willenbrock and D.A. Dicus, Phys. Rev. D **34**, 155 (1986); C.P. Yuan, Phys. Rev. D **41**, 42 (1990); R.K. Ellis and S. Parke, Phys. Rev. D **46**, 3785 (1992); T. Stelzer, Z. Sullivan and S. Willenbrock, Phys. Rev. D **56**, 5919 (1997) hep-ph/9705398.

[32] J. Huston, S. Kuhlmann *et al.*, Phys. Rev. D **58**, 114034 (1998), hep-ph/9801444.

[33] S. Cortese and R. Petronzio, Phys. Lett. B **253**, 494 (1991). T. Stelzer and S. Willenbrock, Phys. Lett. B **357**, 125 (1995), hep-ph/9505433; M. C. Smith and S. Willenbrock, Phys. Rev. D **54**, 6696 (1996), hep-ph/9604223.

[34] A.P. Heinson, A.S. Belyaev and E.E. Boos, Phys. Rev. D **56**, 3114 (1997), hep-ph/9612424; T. M. Tait, Phys. Rev. D **61**, 034001 (2000), hep-ph/9909352; A. Belyaev and E. Boos, Phys. Rev. D **63**, 034012 (2001), hep-ph/0003260; S. Zhu, Phys. Lett. B **524**, 283 (2002) [Erratum-ibid. B **537**, 351 (2002); B. W. Harris, E. Laenen, *et al.*, Phys. Rev. D **66**, 054024 (2002); J. C. Campbell, F. Tramontano, Nucl. Phys. B **726**, 109 (2005), hep-ph/0506289.

[35] D. R. Stump, Prepared for 31st International Conference on High Energy Physics (ICHEP 2002), Amsterdam, The Netherlands, 24-31 Jul 2002. Published in Amsterdam 2002, ICHEP 265-267.

[36] W.J. Stirling, A.D. Martin, R.G. Roberts, R.S. Thorne, AIP Conf.Proc.747:16-21,2005.

[37] W. Hollik, J. I. Illana, S. Rigolin, C. Schappacher and D. Stockinger, Nucl. Phys. B **551**, 3 (1999), hep-ph/9812298.

[38] G. Eilam, J. L. Hewett and A. Soni, Phys. Rev. D **44**, 1473 (1991); B. Mele, S. Petrarca, A. Soddu, Phys. Lett. B **435**, 401 (1998).

[39] C.-H. Chang, X.-Q. Li, J.-X. Wang, M.-Z. Yang, Phys. Lett. B **313**, 389 (1993) C.-S. Huang, X.-H. Wu and S.-H. Zhu, Phys. Lett. B **452**, 143 (1999).

[40] J. A. Aguilar-Saavedra, Phys. Lett. B **502**, 115 (2001); J. A. Aguilar-Saavedra, Acta Phys. Pol. B **35**, 2695 (2004).

[41] H. E. Haber and G. L. Kane, Phys. Rept. **117**, 75 (1985) .

[42] J. F. Gunion, H. E. Haber, G. L. Kane, S. Dawson, *The Higgs Hunters Guide*, (Addison-Wesley, Menlo-Park, 1990).

[43] S. Glashow and S. Weinberg, Phys. Rev. D **15**, 1958 (1977).

[44] S. Bèjar, J. Guasch, J. Solà, Nucl. Phys. B **600**, 21 (2001), hep-ph/0011091.

[45] J. L. Diaz-Cruz *et al.*, Phys. Rev. D **41**, 891 (1990).

[46] A. Antaramian, L. J. Hall and A. Rasin, Phys. Rev. Lett. **69**, 1871 (1992), hep-ph/9206205; L.J. Hall and S. Weinberg, Phys. Rev. D **48**, R979 (1993), hep-ph/9303241; M. Luke and M.J. Savage, Phys. Lett. B **307**, 387 (1993), hep-ph/9303249; W.S. Hou, Phys. Lett. B **296**, 179 (1992).

[47] S. Bar-Shalom, G. Eilam, A. Soni and J. Wudka, Phys. Rev. D **57**, 2957 (1998), hep-ph/9708358; J. L. Diaz-Cruz, M. A. Perez, G. Tavares-Velasco and J. J. Toscano, Phys. Rev. D **60**, 115014 (1999), hep-ph/9903299.

[48] D. Atwood and M. Sher, Phys. Lett. B **411**, 306 (1997), hep-ph/9707229.

[49] G. Eilam, J. L. Hewett and A. Soni, Phys. Rev. D **44**, 1473 (1991); S. Bejar, J. Guasch, J. Sola, (RADCOR 2000) Carmel CA, USA, 11-15 September, 2000, `hep-ph/0101294`.

[50] D. Atwood, L. Reina and A. Soni, Phys. Rev. D **55**, 3156 (1997), `hep-ph/9609279`.

[51] R. A. Diaz, R. Martinez, J-A. Rodriguez, `hep-ph/0103307`.

[52] D. Atwood, L. Reina and A. Soni, Phys. Rev. D **72**, 055018 (2005), `hep-ph/0506167`.

[53] J. L. Diaz Cruz and D. A Lopez Falcon, Phys. Rev. D **61**, 051701 (2000), `hep-ph/9911407`.

[54] C. S. Li, R. J. Oakes and J. M. Yang, Phys. Rev. D **49**, 293 (1994).

[55] G. Couture, C. Hamzaoui and H. Konig, Phys. Rev. D **52**, 1713 (1995).

[56] J. L. Lopez, D. V.Nanopoulos and R. Rangarajan, Phys. Rev. D **56**, 3100 (1997), `hep-ph/9702350`

[57] J. Guasch and J. Solà, Nucl. Phys. B 562, 3 (1999).

[58] J. Guasch and J. Sola, Nucl. Phys. B **562**, 3 (1999), `hep-ph/9906268`; G. Eilam, A. Gemintern, T. Han, J.M. Yang, X. Zhang, Phys. Lett. B **510**, 227 (2001) .

[59] G. M. de Divitiis, R. Petronzio and L. Silvestrini, Nucl. Phys. B **504**, 45 (1997), `hep-ph/9704244`.

[60] J. J. Cao, G. Eilam *et al.*, `hep-ph/0702264v3`, submitted on 26 Feb 2007, last revised 11 Oct 2007.

[61] J. J. Liu, C. Sheng Li, L. L. Yang and L. G. Jin Phys. Lett. B **599**, 92 (2004), `hep-ph/0406155v3`.

[62] J. M. Yang, B.-L. Young and X. Zhang, Phys. Rev. D **58**, 055001 (1998), `hep-ph/9705341`.

[63] A. Cordero-Cid, M. A. Perez, G. Tavares-Velasco, J. J. Toscano, Phys. Rev. D **70**, 074003 (2004), `hep-ph/0407127`.

[64] X.L. Wang *et al.*, Phys. Rev. D **50**, 5781 (1994); G. Lu, C. Yue and J. Huang, J. Phys. G **22**, 305 (1996).

[65] G. Lu, C. Yue and J. Huang, Phys. Rev. D **57**, 1755 (1998).

[66] C. T. Hill, Phys. Lett. B **345**, 483 (1995); K. Lane and E. Eichten, Phys. Lett. B **433**, 96 (1998); G. Cvetic, Rev. Mod. Phys. **71**, 513 (1999).

[67] G. Lu, F. Yin, X. Wang and L. Wan Phys. Rev. D **68**, 015002 (2003); H. Zhang, `0712.0151 [hep-ph]`.

[68] M. Frank and I. Turan, Phys. Rev. D **72**, 035008 (2005).

[69] J. A. Aguilar-Saavedra, B. M. Nobre, Phys. Lett. B **553**, 251 (2003), hep-ph/0210360.

[70] A. Aranda, J. L. Díaz-Cruz, J. Hernández-Sánchez and R. Noriega-Papaqui, Phys. Lett. B, In Press, Uncorrected Proof, Available online 26 October 2007.

[71] B. Grinstein, V. Cirigliano, G. Isidori, M. B. Wise, Nucl. Phys. B **763**, 35 (2007), hep-ph/0608123.

[72] K. Karafasoulis, A. Kyriakis, H. Petrakou and K. Mazumdar, CMS Note 2006/079 (2006).

[73] E. Malkawi and T. Tait, Phys. Rev. D **54**, 5758 (1996), hep-ph/9511337; A. Datta, J.M. Yang, B. Young and X. Zhang, Phys. Rev. D **56**, 3107 (1997), hep-ph/9704257; R.J. Oakes, K. Whisnant, J.M. Yang, B. Young and X. Zhang, Phys. Rev. D **57**, 534 (1998), hep-ph/9707477; T. Han, M. Hosch, K. Whisnant, B. Young and X. Zhang, Phys. Rev. D **58**, 073008 (1998), hep-ph/9806486.

[74] R. Martynez, M. A. Pérez and J. J. Toscano, Phys. Lett. B **340**, 91 (1994); T. Han *et al.*, Phys. Rev. D **55**, 7241 (1997); G. Burdman, M.C. González Garcý and S. F. Novaes, Phys. Rev. D **61**, 114016 (2000).

[75] F. Larios, M. A. Pérez and R. Martynez, Phys. Rev. D **72**, 057504 (2005); R. A. Diaz, R. Martynez and C. E. Sandoval, Eur. Phys. J. C **46**, 403 (2006), 0509194.

[76] [CDF Collaboration], CDF Note 8888.

[77] V. Abazov, *et al.*, [DØ Collaboration], Phys. Rev. Lett. **99**, 191802 (2007), hep-ph/0702005.

[78] Y. P. Gouz and S. R. Slabospitsky, Phys. Lett. B **457**, 177 (1999), hep-ph/9811330.

[79] [ALEPH Collaboration], ALEPH-Conf 99-023; P. Abreu *et al.* [DELHI Collaboration], Phys. Lett. B **446**, 62 (1999), hep-ex/9903072. G. Abbiendi *et al.* [OPAL Collaboration], Phys. Lett. B **521**, 181 (2001).

[80] A. Heister *et al.*, [ALEPH Collab.], Phys. Lett. B **543**, 173 (2002).

[81] J. Abdallah *et al.*, [DELPHI Collab.], Phys. Lett. B **590**, 21 (2004).

[82] P. Achard *et al.*, [L3 Collab.], Phys. Lett. B **549**, 290 (2002).

[83] G. Abbiendi *et al.*, [OPAL Collab.], Phys. Lett. B **521**, 181 (2001).

[84] S. Chekanov *et al.* [ZEUS Collaboration], Phys. Lett. B **559**, 153 (2003).

[85] A. Aktas *et al.* [H1 Collaboration], Eur. Phys. J. C **33**, 9-22 (2004).

[86] Y. P. Gouz and S. R. Slabospitsky, Phys. Lett. B **457**, 177 (1999), hep-ph/9811330.

[87] [ATLAS Collaboration], CERN LHCC 99-14/15 (1999).

[88] E. Richter-Was, D. Froidevaux and L. Poggioli, ATLAS Internal Note ATL-PHYS-98-131, 1998.

[89] L. Chikovani and T. Djobava, hep-ex/0205016.

[90] T. Sjöstrand and M. Bengtsson, Comput. Phys. Commun. **43**, 367 (1987); T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994); T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[91] , S. R. Slabospitsky, L. Sonnenschein, Comput. Phys. Commun. **148**, 87 (2002).

[92] S. Abdullin, A. Khanov, N. Stepanov, CMS TN/94-180.

[93] S. Slabospitsky, talk at *Physics at LHC 2004*, Vienna, Austria, 13-17 July 2004.

[94] [CMS Collaboration], Physics TDR Vol.I, Chapter 2; http://cmsdoc.cern.ch/ORCA/ .

[95] [CMS Collaboration], J. Phys. G: Nucl. Part. Phys., **34**, 995 (2007).

[96] L. Benucci, A. Kyriakis, *et al.*, CMS Note 2006/093 (2006).

[97] J. Carvalho, N. Castro, *et al.*, Eur. Phys. J. C **52**, 999 (2007), 0712.1127v1 [hep-ex].

[98] S. Abdullin, G. Azuelos, A. Ball *et al.*, (2002), hep-ph/0204087.

[99] T. L. S. Group, CERN-AC-95-05 (1995), hep-ph/0601012.

[100] [CMS Collaboration], CERN/LHCC 1992-3 (1992); CERN/LHCC 94-38 (1994).

[101] [CMS Collaboration], CERN/LHCC 2000-38 (2000); [CMS Collaboration], CERN/LHCC 2002-26 (2002), http://cmsdoc.cern.ch/cms/TDR/DAQ/daq.html.

[102] [CMS Collaboration], CERN/LHCC/2006-001, CMS TDR 8.1, 2 February 2006; CERN/LHCC/2007-009, CMS TDR 8.2-Add1, 5 March 2007.

[103] E. Meschi *et al.*, CMS Note 2001/034 (2001).

[104] M. Pieri *et al.*, CMS Note 2006/007 (2006).

[105] N. Marinelli, CMS Note 2006/005 (2006).

[106] R. Frühwirth and T. Speer, Nucl. Instrum. and Methods, A **534**, 217 (2004).

[107] R. Frühwirth, Nucl. Instrum. and Methods, A **262**, 444 (1987) .

[108] S. Baffioni *et al.*, CMS Note 2006/040 (2006).

[109] E. James, Y. Maravin and N. Neumeister, CMS Note 2006/010 (2006).

[110] S. V. Chekanov, hep-ph/0211298.

[111] G. C. Blazey *et al.*, hep-ex/0005012.

[112] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 31603166, (1993), hep-ph/9305266; S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. B **406**, 187 (1993); M. H. Seymour, Nucl. Phys. **B421**, 545 (1994); J. M. Butterworth, J. P. Couchman, B. E. Cox, and B. M. Waugh, Comput. Phys. Commun. **153**, 8596 (2003), hep-ph/0210022.

[113] A. Heister, O. Kodolova, V. Konoplianikov, S. Petrushanko, J. Rohlf, C. Tully and A. Ulyanov, CMS Note 2006/036 (2006).

[114] V. F. Konoplyanikov *et al.*, Phys. Part. Nucl. Lett. **2** 45 (2005) .

[115] S. Kunori, R. Kinnunen and A. Nikitenko, CMS Note 2001/040 (2001); H. Pi *et al.*, CMS Note 2006/035 (2006).

[116] A. Rizzi, F. Palla and G. Segneri, CMS Note 2006/019 (2006).

[117] C. Weiser, CMS Note 2006/014 (2006).

[118] A. Bocci, P. Demin *et al.*, CMS Note 2006/043 (2006).

[119] V. Innocente, L. Silvestris and D. Stickland, Comp. Phys. Comm. **140**, 3144 (2001).

[120] [CMS Collaboration], http://cmsdoc.cern.ch/OSCAR/.

[121] [GEANT4 Collaboration], S. Agostinelli *et al.*, Nucl. Instrum. and Methods, A **506**, 250 (2003).

[122] Informations available at http://root.cern.ch/.

[123] [CMS Collaboration], Physics TDR Vol.I, Chapter 2; http://cmsdoc.cern.ch/FAMOS/.

[124] J. Campbell, R.K. Ellis, F. Tramontano, Phys. Rev. D **70**, 094012 (2004), http://mcfm.fnal.gov.

[125] J. Campbell, R. K. Ellis, D. Rainwater, Phys. Rev. D **68**, 094021 (2003), hep-ph/0308195.

[126] M. Chiorboli, M. Galanti, A. Tricomi, CMS NOTE 2006/133 (2006).

[127] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07**, 1 (2003), hep-ph/0206293.

[128] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999), hep-ph/9905386.

[129] C. Zecher, T. Matsuura and J. J. Van der Bij, Zeit. fur Phys. C **64**, 219 (1994), hep-ph/9404295.

[130] *CMS values for the (W+- jet) production cross sections*, http://cmsdoc.cern.ch/cms/PRS/gentools/www/xsec/det_w_jet.html.

[131] A. Pukhov *et al.*, hep-ph/9908288.

[132] B. Andersson *et al.*, Z. Phys. C **20**, 317 (1983).

[133] C. Peterson *et al.*, Phys. Rev. D **27**, 105 (1983).

[134] J. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, Phys. Rev. D **69**, 074021 (2004), hep-ph/0312024v2; J. Campbell, R. K. Ellis, F. Maltoni and S. Willenbrock, Phys. Rev. D **73**, 054007 (2006), hep-ph/0510362v2.

[135] W. Beenakker, S. Dittmaier *et al.*, Nucl. Phys. B**653**, 151 (2003), hep-ph/0211352.

[136] D. Benedetti, S. Cucciarelli, *et al.*, CMS Note 2006/119 (2006).

[137] http://cmsdoc.cern.ch/cms/PRS/gentools/www/systematics/syserr.html.

[138] J. D'Hondt, S. Lowette, S. Kasselmann, CMS NOTE 2006/025 (2006).

[139] S. Lowette, J. D'Hondt, J. Heyninck, CMS Note 2006/013 (2006).

[140] P. Bartalini, R. Chierici, A. De Roeck, CMS Note 2005/013 (2005).

[141] M. R. Whalley, D. Bourilkov, and R. C. Group, hep-ph/0508110.

[142] A. Kiryunin, Meeting at MPI on "HEC cold electronic", 16 April, 2007.

[143] R. D. Cousins and V. L. Highland, Nucl. Instr. Meth. A **320**, 331 (1992); G. D'Agostini and M. Raso, CERN-EP/2000-026, 1 February, 2000, hep-ex/0002056; S. I. Bityukov and N. V. Krasnikov, Modern Physics Letters A **13**, 3235 (1998); I. Narsky, Nucl. Instr. Meth. A **450**, 444 (2000).

[144] S. I. Bityukov and N. V. Krasnikov, Nucl. Instrum. and Methods, A **452**, 518 (2000).

[145] http://cmsdoc.cern.ch/~bityukov.

[146] G. Punzi, hep-ph/0308063v2.

[147] M. Veltman, Nucl. Phys. B **123**, 89 (1977).

[148] [LEP Higgs Working Group], Phys. Lett. B **565**, 61 (2003), hep-ex/0306033.

[149] [CMS Collaboration], CERN/LHCC 97-032 (1997).

[150] [CMS Collaboration], CERN/LHCC 97-033 (1997); [CMS Collaboration], CMS TDR 4, Addendum CERN/LHCC 2002-027.

[151] [CMS Collaboration], CERN/LHCC 97-031 (1997).

[152] L. Borrello, A. Messineo, E. Focardi, and A. Macchiolo, CMS Note 2003/020 (2002).

[153] [TOTEM Collaboration], CERN/LHCC 2004-002 (2004); [TOTEM Collaboration], CERN/LHCC 2004-020 (2004).